Student-centered instruction and math and science achievement in the post-Soviet state

A mixed methods analysis

Carina Omoeva

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2012

ABSTRACT

Student-centered instruction and student achievement in the post-Soviet State

Carina Omoeva

This dissertation examines the relationship between the use of student-centered instructional methods in teaching mathematics and science, and achievement of fourth graders in these subjects. The context for analyzing this relationship is the post-Soviet region, with Kazakhstan selected as the main focus of the study, and the object of its in-depth qualitative case analysis. The measures of student-centered instructional methods are drawn from student surveys administered as part of the 2007 TIMSS study, while the TIMSS test scores serve as measures of student achievement. The quantitative analysis finds that student-centered instructional methods generally show no statistically significant relationship with student achievement across all country datasets included in this analysis. The qualitative case study follows up on these results in the context of Kazakhstan, and finds that while student-centered instruction is hailed as the pathway to reform of education in this post-Soviet country, the state lacks the capacity to engage in instructional reform and improvement of teacher quality, while the teachers look to the state to guide them in choosing the right instructional methods. As a result, instructional transformations take place in highly haphazard, heterogeneous ways, while teachers require direction, guidance, and support from the central state to effectively implement student-centered methods in their lessons.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

ACKNOWLEDGEMENTS

I want to thank my advisor, Gita Steiner-Khamsi, for her guidance and feedback throughout this dissertation project, and for helping me build a broader, systemic, and policy view on what began as a micro-level research interest. I am also deeply grateful to my professors and dissertation committee members Henry M. Levin and Douglas Ready for the help and guidance they provided to me during the difficult phases of data analysis and interpretation. Many thanks to committee members Iveta Silova and Timothy Frye, for their thoughtful comments on this dissertation and their ideas for further in-depth research.

I am thankful to my wonderful friends at Teachers College, Columbia University – Andrew Shiotani, Christine Harris-Van Keuren, Erin Weeks-Earp, and Steven Ehrenberg, for their valuable feedback at all stages of this dissertation, helping me think through it from the initial concept to its final form, serving as the sounding board for my ideas and the very first readers of my drafts, and always being there to offer a word of encouragement. David Sprague, for stimulating discussions about instructional reform and student achievement, and his comments on the early write-ups of my results chapters. I owe huge thanks to Hanymkyz, Khanen, and Gulnaz Muratbekova for their invaluable help in setting up interviews during the qualitative data collection phase in Kazakhstan. The Embassy of Kazakhstan in Washington, DC, was also gracious in helping to arrange contact with the Ministry of Education in Astana, Kazakhstan, and Ayman Kashkeyeva at the Ministry of Education was helpful in arranging meetings and obtaining policy documents.

Finally, my biggest debt of gratitude is to my family: my husband Timur, on whose help in anything related to the production of this dissertation I could (and did) always count, and my daughters, Tina and Bela, for their loving encouragement, unending support,

and positive endurance throughout the countless days and nights devoted to my research, analysis, and dissertation writing. Murad, Liliya, and Asylbek Omoev, for their support at every step of the way, making this project possible to accomplish. My mother, Medira Irgebayeva, for her support and encouragement, and for setting an example, long ago, of how far one must go in one's academic pursuits.

CHAPTER 1.
PROBLEM STATEMENT

1.1. INTRODUCTION

What is the best way to learn math and science? Is it best learned through direct instruction by the teacher, memorization of core principles, and long hours of practice in solving a variety of math or science problems, until a level of high proficiency is reached, and the school curriculum is conquered? Or, should the emphasis be placed on the process itself, with the student learning math and science through interaction, experimentation, observation, group work, and discussion? The question of how people learn lies at the heart of pedagogy as science, and to date there are no definitive answers.

While different teaching methodologies can – and by all means, should – be combined depending on the task at hand and the learning needs of the students, it has become common to broadly group the former methods as "traditional instruction" and the latter "student-centered instruction", as a way of noting the emphasis of the teaching and learning dynamic in the classroom. This juxtaposition sometimes also carries an implicit judgment: student-centered instruction is seen as the more "modern" way of teaching, more in-tune with the changing requirements of the modern economy and the democratic social landscape that prevails in developed countries (see Literature Review in Chapter 3). Traditional instruction, by contrast, is viewed increasingly as a thing of the past, particularly in countries where states seek to redefine their identities on the world stage. However, is this conceptual shift justifiable from the point of view of improving student achievement? And if so, are there certain parameters of the system that make it possible to attain success with one approach to instruction, but make it less certain with another? These are the guiding questions that prompted my interest in researching the relationship between

instruction and achievement in a post-Soviet environment, and led to the writing of this dissertation. It is my hope that through a combination of quantitative analysis and a qualitative case study, I make a contribution to the understanding of the relationship of student-centered instruction and achievement as that which is far from positive or causally deterministic, which requires substantial capacity and commitment on the part of the state to be effective – even in a system where the central state has all of the mechanisms to enforce compliance with its policies.

1.1.1. OVERVIEW

This dissertation focuses on the evaluation of the role of student-centered instructional methods in student achievement in math and science, in the context of the post-Soviet state. I first examine whether or not there is a direct relationship between student-centered teaching practices and achievement outcomes, through an analysis of the TIMSS 2007[1] student achievement data from post-Soviet countries, with an additional focus and extended analysis of the data on Kazakhstan. The results of my quantitative analysis are then taken a step further, with qualitative exploration of the context of the post-Soviet education system. Using the case of Kazakhstan, I examine what challenges and opportunities face a highly centralized state as it engages in education reform, what mechanisms it employs to achieve its goals, and particularly, how its actions are reflected in the realities of schools and teachers.

[1] TIMSS: Trends in International Math and Science Study is a regular international assessment of the math and science achievement outcomes of students in 4th and 8th grades of school, established in 1995 and repeated every four years, with the most recent (at the time of this dissertation) completed in 2011.

I use this contextual information to build an extension of the argument advanced by Carnoy et al. (2007) on the role that the state can play in building a foundation for education quality, and the mechanisms that are available to a centralized state with a tightly coupled system. I argue that while such mechanisms do indeed facilitate faster implementation of reform and help ensure that elements of the system fulfill their role in achieving its goals, such systems create an over-dependence on the center and are unable to successfully assimilate innovative practices from within, without an explicit roll out from the state. Furthermore, it is important to distinguish between the different types of leadership that can be exercised by a centralized state: administrative, financial, and instructional, and it is this lattermost category of leadership that must be present in order for a teaching quality improvement to take place in a nation's classrooms.

As part of the quantitative analysis, the relationship between instruction and achievement is examined in this dissertation both in cross-country perspective, by looking at data from several post-Soviet states at the same time and interpreting their similarities and differences in results; and in more detail, focusing on Kazakhstan's data, where even greater care is placed on bias control through techniques such as matching and weighting on the propensity scores. In the qualitative case study, I build context through interviews with educators in Kazakhstan at different levels of the system, as well as policy documents reflecting the vision and narrative of the state. Both of these elements – the quantitative and the qualitative – provide two angles on the issue of student-centered instruction and its place in pedagogy, with respect to measurable student achievement outcomes. As a result, I not only provide an examination of the contribution of student-centered instruction to student achievement in math and science in such a context, but also offer an explanation of my findings, pointing to the context of the education system, the realities of the classroom,

the challenges and incentives facing teachers, and other structural and cultural traits that are still vividly present in the post-Communist society, and particularly, in the relationship between the state and the individual in the education system.

1.1.2. QUALITY OF LEARNING = STUDENT-CENTERED TEACHING?

While few people would doubt the importance of good teaching quality for achieving meaningful results in education, it would be fair to say that nothing has generated more debate and less agreement among scholars and practitioners than the definition of what constitutes good teaching. Vast amounts of literature have been produced on the subject of teacher quality and its various aspects, from teacher training and experience in the classroom to compensation and performance incentives. However, capturing the effects of each of these features has proven difficult: only small or negligible effects of these measurable traits have been detected by large-scale studies, and none have consistently shown a strong effect on performance. Evidence of the effect of instructional methods has been even more problematic, due to the highly arbitrary nature of decisions involved in building instructional models, and the general lack of a truly randomized setting where any particular set of techniques would be tested.

Nonetheless, since the early 20th century, when the Progressives in the West and scholars such as Lev Vygotsky in the newly established Soviet Union advanced their theories of child development and socialization, the calls for a more individualized, child-centered approach to teaching have not seized. During the Soviet era in Eurasia, despite the presence and recognition of the early works of Vygotsky, Sukhomlinsky and many other educators in innovative instruction, student-centered methods did not enter mainstream schools. However, they are increasingly gaining greater ground post-independence, as they

are promoted by international donors and educational NGOs in countries where traditional methods prevailed throughout decades. In many ways, the shift towards student-centered instruction may be viewed as an example of an adoption (albeit, as I will show, mainly discursive) of a "global policy" due to its perceived, rather than demonstrated, benefits.

There are different theoretical arguments explaining the phenomenon of policy adoption. One view, in line with the canons of neoinstitutionalism, is that it constitutes an example of a burgeoning "world culture", a worldwide dissemination of Western educational norms and practices. As Baker and LeTendre (2005) advance, the widespread understanding communicated through international organizations results in an acceptance of ideas and leads to similar meanings, all of which happens in "soft, almost imperceptible ways" (p.10), such that education policy makers have little control over the influences on their education systems. If student-centered instruction is the way that schooling as a "global institution" is organized, then national education systems are likely to feel the pressure to conform to this global trend.

Another view would hold, however, that "traveling" reforms are, in fact, conscious choices made by policy makers in order to advance their domestic agendas, or to project a certain image to the international audiences (see, for example, Steiner-Khamsi & Stolpe, 2006; or Steiner-Khamsi et al., 2005). States adopt reforms on a discursive level, without actually implementing the changes that it would take to transform their national institutions. This engagement in "double-talk" is also conscious, and takes place for rational reasons: policy makers receive credit for adopting the language of modernity, but suffer little political risk by not carrying out serious structural reforms, the outcomes of which are less clear than the payoff from "double-talk". Governments may also choose to engage in double talk in order to preserve the status quo, while creating the appearance of reform.

Regardless of the theoretical explanation driving them, the presence of such reforms is evident. In the policy rhetoric of reform, student-centered instruction appears to be equated with modernization and development, and is conceptually connected with the idea of increased democratization in the education system. The new National Education Strategies of post-Soviet states, for example, explicitly state that a new approach to teaching would be incorporated into the education reform package, as it is seen as a way of improving the preparedness of a nation's youth for the jobs of the future, and hence, raising its competitiveness on the global labor market (e.g. Government of Russian Federation, 2009; Government of Kazakhstan, 2009). However, when it comes to understanding the links between student-centered pedagogy and achievement outcomes, there is a lack of rigorous research documenting this relationship, not just in the region but in literature as a whole. Evidence mostly comes from case studies where specific programs and practices are evaluated in depth, but which provide few generalizable lessons learned and replicable models for the general population (Darling-Hammond et al., 2008; Johnson, Johnson, and Stanne, 2000), or quantitative studies that use observational data from international achievement studies and fail to correct for the highly nonrandom nature of the sample, undermining the validity of the causal analysis. In the former Communist states, calls for democratization of teaching methods began long before the breakup of the Communist system itself (Webber, 2000), and were made mostly in a normative way (such as the value of changing the schooling culture per se), rather than as a way of boosting achievement outcomes. In recent studies, capturing large-scale effects of interactive pedagogy has been problematic, with mostly no significant findings, or small but negative coefficients (House, 2002; OECD, 2008; Aypay, Erdogan, and Sozer, 2007)

Notwithstanding the value of student-centered, active learning methodology in and of itself, in a world where emphasis is increasingly being put on cross-national comparison, and governments look for the most effective ways to bring about higher achievement outcomes on a large scale, it is a legitimate question to ask whether removing rigidity in teaching will actually bring about better learning. In view of these broader implications for education policy, my dissertation seeks to make an addition to the growing body of research that uses large-scale datasets to explore the correlations between practice and achievement outcomes in a "real world" setting, rather than in a laboratory environment.

1.1.3. ACHIEVEMENT OUTCOMES: TAKING A SYSTEMIC VIEW

Any post-Soviet state bears the common characteristics in its education system: universal compulsory enrollment, a comprehensive approach to curriculum with vertical and horizontal alignment of subject matter and nationally universal content, a tradition of equitable levels of inputs across public schools, and a largely female teacher cadre. The traditional approach to instruction is also often seen as an inherent feature of the post-Soviet school, despite the longtime presence of student-centered instructional models (see Chapter 2 and 3). Since the breakup of the Soviet Union, its former republics became the recipients of substantial Western and international donor assistance, which included involvement with reform efforts in all sectors of the economy and in many cases, loan or grant funding for infrastructure projects and long-term development needs. In education, non-state actors such as international education NGOs have been most active, building networks of schools and establishing professional development opportunities for teachers, providing school supplies and funding small-scale school improvement projects. The weakness of government capacity in post-Soviet countries to implement reforms prompted some of the NGOs, such as Open Society Institute, to work entirely at the school level with

minimal involvement, the local or national governments. Larger donors, as well as bilateral ones, invested their effort heavily into building capacity in the state administrative structures – often wasting their resources to personnel turnover fueled by political instability. The question of State capacity continues to be salient as governments seek a greater role in the implementation of donor-funded projects, following the Paris Declaration on Aid Effectiveness (2005).

The availability of internationally comparable data on student achievement often prompts immediate causal inferences about national education systems: governments in top performing countries are usually given credit for the successes on international tests, often without questions about the broader contextual factors affecting student achievement. Given the heterogeneity of student populations across nation-states, and the diversity of education structures and policies, exploring how the effectiveness of specific teaching practices may be affected by broader environmental factors is a fascinating undertaking, and the availability of a common metric measuring outcomes makes it possible to quantify these relationships, albeit still bound by the limitations of the available metrics. In this dissertation, I will explore how the findings on the links between student-centered teaching practices and achievement may be explained by the capacity of State to follow through on its vision of improving instruction, and offer a cautionary argument against tying achievement outcomes directly to education policies or other replicable elements of the education system.

A large body of literature connects the ability of an education system to produce quality learning for its students to the capacity of the State to deliver its services, and one such argument was used to set up the theoretical framework of this dissertation (Carnoy et al, 2007). However, much like the uncertainty surrounding the effect of student-centered

instruction on achievement, there is no firm agreement among scholars and the public at large on what the role of the State should be vis-a-vis the individual in education. In light of this uncertainty, one may argue that the changes that took place in the national education systems in the post-Soviet states may be taken as manifestations of the predominant philosophy on the role of the State in education, and hence, as potential lenses through which to interpret the results on the achievement outcomes. Therefore, as I examine the context of the post-Soviet Kazakhstan, a highly centralized system with a high concentration of power at the top executive level, I start with the official state rhetoric as a demonstration of its vision for educational development, and compare this vision to the realities of the teachers and schools, gauged through my qualitative data collection. I apply a theoretical framework built on an argument in favor of a great state role in education, developed by Carnoy et al. (2007) to examine whether the role of the state in Kazakhstan's educational development has been positive, and more importantly – whether any evidence exists that it has positively affected the instructional environment in the nation's classrooms, or taken any steps to improve the quality and the mode of teaching.

The context selected for this study will make it especially interesting to examine the relationship between the State and education: the post-Soviet region as a whole, and Kazakhstan in particular, experienced profound political, economic, and social changes over the two decades following independence. I make no attempt here to fully explain the variance in achievement outcomes across post-Soviet states. Instead, I offer a brief look at the heterogeneity of relationships between a defined set of instructional practices and student achievement across countries. As I go into greater depth with the data from Kazakhstan, I explore how the choice of pedagogy is influenced by the broader capacity of

the institutions of the centralized post-Soviet state, and attempt to at least partially explain the results of quantitative analysis in Kazakhstan.

The theoretical framework and the methods for this dissertation are designed to address the research questions that reflect both my interest in examining the relationship between student-centered instruction and achievement per se, as well as in understanding the role of a centralized state in fostering its successful implementation.

1.2. RESEARCH QUESTIONS

This dissertation is a mixed methods study, where the qualitative component serves as an extension of the quantitative analysis. The research questions that I tackle are listed below.

1. What is the relationship between student-centered instruction and student achievement in math and science in Kazakhstan?
    a. What is the association between student-centered instructional methods and achievement in math and science in Kazakhstan? How does it compare with that in the other post-Soviet states?
    b. What types of students and teachers are more likely to have been studying and teaching in student-centered environments?
2. What are the systemic and contextual factors affecting the use of student-centered methods in classroom instruction in math and science?
3. What role does the state play in the post-Soviet transformations taking place in education in Kazakhstan?

While there is no strict delineation of methods for each of the research questions, RQ 1 is mostly addressed through quantitative analysis, and RQ's 2 and 3 – through the

qualitative case study. More detail on the methods for addressing these research questions is presented in Chapter 5.

1.3. CONTRIBUTION TO THE FIELD

As I mentioned above, generalizable evidence of the effect of specific types of instruction on students has been difficult to observe on a large scale. Studies that used international achievement data have generally found either no effect of student-centered instruction on achievement outcomes, or a negative effect in some settings (OECD 2008; House 2005). My earlier work with the TIMSS dataset (in Silova 2011) indicated that the relationship between instruction and achievement in math and science is complex, nonlinear, and heterogeneous across the countries included in the analysis. With this study, I intend to move the field closer to causality between instruction and achievement that earlier works have not been able to do, by virtue of using a more complex statistical technique, such as matched sampling with propensity scores, and approaching the data differently to create a more meaningful coding of "treatment", or student-centered instruction. I hope that the findings from this study can enrich the literature on student achievement by providing a more rigorous estimate of what, if any, relationship exists between exposure to student-centered teaching and achievement.

Since I arrive at my empirical findings using matched sampling methods that have been established over two decades ago and are now increasingly gaining ground as a method for causal research (Rosenbaum and Rubin, 1984; Rubin, 1997; Hill et. al, 1999; D'Agostino et al., 2000; Rubin, 2001), this dissertation also seeks to make a contribution to the methodology for understanding the relationship between student-centered instruction and achievement. The study will make a case for greater use of quantitative analysis on par

with in-depth qualitative studies, which have dominated the sub-field of teaching methods and student-centered instruction. In addition, since this study is not an evaluation of a set program or policy, but rather an exploration of the impact of existing teaching practices, as part of my methods I will make a series of decisions about how the "treatment", or student-centered instruction, should be coded, and what combination of practices would be considered full treatment. In this respect, this dissertation will seek to operationalize the notion of student-centered instruction, such that the same principle and same method for a treatment definition could be further applied in analyses of data from subsequent rounds of TIMSS or from comparable international and national achievement studies.

Finally, this dissertation will extend and challenge the theoretical argument of Carnoy et al. (2007), by offering a case of a post-Soviet state, where the link between the government's vision of educational development and its practical implementation is not as tight as one would expect in an environment with such a high concentration of power and resources in its central bodies. I will also show that it is important to consider not only the strength or the tightness of control mechanisms, but distinguish between the types of leadership and the belief system regulating the dynamics of teaching and classroom management. I will offer evidence of heterogeneity and complexity surrounding instructional choices and the quality of teaching in Kazakhstan's schools, and suggest a new interpretation of the role of centralized states in education, with a greater recognition of both the advantages of extended presence of the state in all matters of public life, as well as the threats it potentially poses to the quality of education, should the capacity of central agencies to lead schools towards continuous improvement gradually (or suddenly) deteriorate.

1.4. STRUCTURE OF THIS DISSERTATION

This dissertation is structured as follows. Chapter 2 provides relevant historical and contextual background on the post-Soviet educational environment, the history of student-centered instruction in the region, and the more recent developments in education policy in Kazakhstan. Chapter 3 follows this introduction with a review of the literature on student-centered instruction, both in theory and in practice, as well as a review of the cases of high-achieving states, seen through the lens of the role of the state. Chapter 4 builds the theoretical framework based on these two strands of literature. Chapter 5 outlines the methods used in this dissertation in order to address the research questions noted above. Chapter 6 and Chapter 7 present the results of quantitative and qualitative analyses, respectively. Chapter 8 concludes and presents a potential agenda for future research.

CHAPTER 2.
BACKGROUND

In this chapter, I offer a brief overview of the context of the post-Soviet states, starting with the overall commonalities between the seven states included in the quantitative analysis, followed by more detail on the current context in Kazakhstan. The purpose of this chapter is to familiarize the reader with the education systems in the region, which serves as a background for the understanding of the quantitative findings, as well as the qualitative analysis of Kazakhstan. This chapter also offers a glimpse of history of education in the region, and marks the most notable developments, including the choices of instructional approach across the education system, as well as the various contributions to student-centered instruction and other instructional innovations prior to the modern period of the post-Soviet state.

2.1. THE "TRANSITIONAL STATES"?

As will be described in Chapter 5, the quantitative part of this dissertation begins with a regression analysis of the TIMSS 2007 data from seven former Soviet republics in Eurasia. The countries that were selected for this analysis share a common history as republics of the former Soviet Union, a once-mighty super-state whose achievements in science at the height of the Cold War served as an impetus for an overhaul of the U.S. education system, but which has since fallen apart, causing a decade-long crisis in all sectors of public life in its fifteen republics. Some of them, such as the Baltic States, took less time to rebuild their economies and have become exemplars of democracy in Eurasia, eventually joining the European Union. Others, such as Russia and Kazakhstan, remarkably turned around their economic fortunes, albeit with little success on the democratic front; still others, like Georgia, are struggling to meet the basic needs of their citizens. Given the

profound changes in the political systems, economies, and social sectors among the former Soviet republics, a legitimate question arises whether they still belong in the same group - the "transitional states" - as they are frequently referred to in the press and donor talk. To what extent the history of the Soviet rule still holds them together is a question that political scientists have posed repeatedly in the last nearly two decades.

Among donors, the rule appears to be to group the post-Soviet republics with Eastern European states, and in the case of World Bank, with Europe as a whole (www.worldbank.org/eca). The U.S. State Department recently reorganized its regional bureaus such that some of the post-Soviet republics were shifted into the South Asia region. Indeed, the distance between some of the members of the demised Soviet Union, not just geographically, but also in terms of their economic development, seems too vast for them to belong in the same geopolitical grouping, as some have joined the European Union while others are struggling to make ends meet for much of their population. Throughout the first decade of independence, various efforts aimed at securing regional cooperation and foster interdependence among the new formed CIS members have been disappointingly ineffective, with regulations stalling trade rather that facilitating it (Libman and Vinokurov, 2010). CIS members have looked to outside partners for support throughout their periods of economic reform, as witnessed by the firm presence of international donors in the region (Dabrowski, 1995; Davis & Dombrowski, 2000), and their involvement in setting and implementing the reform agenda. Russia's continued dominance as a great power in the region, however, is hardly doubted, especially apparent when it becomes contested and reaffirmed, as in the case of color revolutions in Georgia and Ukraine, or armed border conflicts as in the brief Russia-Georgia war of 2008.

In sum, the level of inter-state connections among the post-Soviet republics has by all means greatly diminished, voluntarily or not, since the days of the USSR. However, as Libman and Vinokurov (2010) argue, the region can still be viewed as possessing strong economic and cultural commonalities. In fact, the economic revival of Russia, Kazakhstan, and perhaps to a lesser extent, other post-Soviet republics in the early 2000's reaffirmed their integration trends, as trade volumes grew and labor and education migration increased exponentially (Ibid.) Furthermore, greater resources also propelled Russia's ambitions to reclaim its "near abroad" as a sphere of firm influence. Reform patterns across country-members CIS often cross-reference efforts taking place in other post-Soviet states as a way of added validation with domestic audiences that may otherwise be skeptical of the changes. In system theory, exemplified most notably by the works of Schriewer (1990), Steiner-Khamsi (2004), Schriewer and Martinez (2004), and Steiner-Khamsi and Stolpe (2006), such processes are called externalization: actors in self-referential systems may adopt reforms from outside, in order to promote their political or economic agenda, or "sell" contestable policies of preference domestically. True, these countries have gone their separate ways, but as they did, they still look to each other for reform ideas and lessons learned, at the same time as they reestablished priorities and reform agendas across all sectors.

The education systems of these republics, these essential elements of the social sector of immense importance to society, have undergone changes in the past two decades as countries sought to redefine their national identities and purge symbols of the past ideology. All of the countries have adopted national education strategies, in which they proclaim adherence to the highest standards of achievement, development of key skills and competencies, and support for innovation (see, for example, national strategies referenced

in Kovaleva and Krasnyanskaya, 2007; Mkrtchian, 2007; Government of Kazakhstan, 2008; etc). As we assess the commitment of the post-Soviet states to advance positive changes in the education sector, it may be useful to glance over their journey from the centralized Soviet system to their current state. I will use this section to provide a brief overview of that common starting point, as well as of the key challenges that countries faced in the past two decades.

2.1.1. THE SOVIET PERIOD AND EARLY REFORM EFFORTS

While some states had established public schools prior to the Soviet period, these schooling networks were subsumed into the unified system with a single curriculum, single textbooks, standards, examinations, and teacher preparation institutions, as part of the general socialist path of development. Universal access to free public education until the age of fifteen was guaranteed by the State as a means of equalizing educational opportunity and minimizing social stratification. Compulsory education in comprehensive basic schools was nine years, beyond which the students either continued to the remaining two grades of upper secondary school and graduate, or transfer into a vocational school for two years before graduating with a secondary specialized degree.

The Communist belief in equal abilities of all children regardless of their families, and rejection of the theories of individual differences, resulted in educational policies that offered equal educational services without any differentiation in the curricula or teaching methods across the Soviet Union (Fugueroa, 1963; Shimoniak, 1970) – at a time when innate differences were widely accepted in the West, and countries such as the United Kingdom organized the entire educational process around the perceived intrinsic intellectual differences among children. Thus, the Soviet educational project was inherently

progressive, and aimed at expanding educational opportunity to all citizens – as part of its official ideology (a similar ideology is reflected in the actions of the Cuban state, referenced in Chapter 3 and Chapter 4). In practice, however, specialized schools for the gifted in specific disciplines, such as the sciences or languages, did exist – and in fact, flourished – serving the need for differentiation and efficiency in producing quality cadre for the society, at the same time providing highly specialized labor for the industrial and military complex. While basic compulsory education was mostly unified and egalitarian, greater social stratification began in the post-compulsory stage of public education, with the more academically able continuing to attain a higher education diploma, while the less able gained vocational certificates (Shimoniak, 1970).

Perhaps the most important feature of the Soviet education system was its emphasis on the collective over the individual, following Makarenko's theories of the collective spirit and the importance of self-discipline and conformity with group interests and preferences (Bowen, 1962; Shimoniak, 1970). Shimoniak (1970) noted with curiosity that the competitive activities within the schools were carried out between groups rather than between individuals. One may note, however, that it is a feature that is gaining ground in Western schools now, with a growing emphasis on teamwork and group work. Much has been written over the years on the rigidity of the Soviet school as an ideological apparatus, with its pervasive indoctrination of youth with Communist values and a repression of innovation and independent thought (Turkevich, 1949; Shimoniak, 1970; Reilly, 1996). And while the groundbreaking works of Lev Vygotski in the 1920's served to establish educational psychology as a separate discipline, little was done over the years to incorporate his theories into the instructional practices of the millions of Soviet teachers. Educational psychology remained largely a field of theorizing, and less of practical

application, despite the high regards from Western social scientists for the Soviet educational psychology as "extremely ingenious and generally profound" (O'Connor, 1963, p.51). As much as the Soviet educators of the early years were impressed and influenced by Western philosophers such as John Dewey, methods and practices developed with democratic goals in mind were deemed unfit for the Soviet society, and denounced as bourgeois ideas detrimental to the advancement towards Communism (Shimoniak, 1970). The emphasis on the belonging to the State lent itself to teaching methods that placed the teacher entirely in the center of the educational process, with each student and her parents expected to keep pace with the comprehensive curriculum and standards as part of their responsibility as Soviet citizens. Parents often deferred matters of child upbringing to the schools, and could receive visits from school staff to their homes to ensure that sufficient attention and emphasis is placed on studies within the household (Ibid., pp 208-209; Waddington, 1963). This dynamic is still present in Cuba, where the school is held accountable for student progress, and may put pressure on families to ensure that their child has the necessary conditions to fulfill his academic requirements (Carnoy et al. 2007).

As the Soviet economy suffered the consequences of arms race and imperfect management, as well as the decline in oil production and exports, by the 1980's its gigantic education system began to suffer resource shortages, both in the infrastructure and in human resources. Reilly (1996) argues that the economic failures of the Soviet State were the impetus for political and social reforms, and particularly, for *perestroika* and *glasnost*. While political scientists still argue about the direction of causality between the economy and the political environment when it comes to the breakup of the once-mighty empire (Kalyvas, 1999), numerous accounts and documents attest to the fact that reforms of the education system were initiated during the Soviet period. Importantly, these calls for

reform in education focused first and foremost not on efficiency, accountability, or achievement gaps, as is often the case in the West, but on *redefining the relationship between the State and the individual* in the learning process. Progressively-minded educators in the Soviet Union sharply criticized the state educational machine as an impersonal, dull, and ineffective bureaucracy that was not capable of developing the intellectual and creative potential of the country's youth, and offered no reward to talented teachers that stepped out of the boundaries of traditional pedagogy (Reilly, 1996; Webber, 2000; Eklof et al., 2000). In particular, the Academy of Pedagogical Sciences was starkly criticized in the media in the late 1980's as inert, removed from the reality of the classroom, and unable to meet the needs of the teachers (Reilly, 1996).

However, it would be wrong to assume that no pedagogical innovation took place until the breakup of the Soviet Union, or that innovation is intrinsically linked with democratization and post-independence. In fact, many of the models of student-centered instruction based on the works of Soviet and Russian educators received notice and stirred great interest among teachers well before the 1990's. As the country's leadership embarked on a profound overhaul of the economy in the *perestroika* process, reform of education received additional impetus as an integral part and a precondition for the success of perestroika itself (Reilly, 1996). Webber (2000) notes the popularity in the mid-1980's of the *Pedagogy of Cooperation*, advanced by a number of innovative educators independently of the Academy of Pedagogical Sciences. The innovative ideas gained substantial press coverage in the main education outlets, including *Uchitelskaya Gazeta* (Teachers' Newspaper), and led to the formation of professional networks of teachers, such as the Creative Union of Teachers. However, as the author notes, the innovations were too often lacking suggestions for practical implementation and dependent on the personal charisma

of the teachers, which was not sufficient for incorporation into the mainstream teacher training. Furthermore, even when innovative ideas were incorporated in a strategy for reform, a multitude of reasons, including political instability, inertia of the state bureaucracy, and lack of funding, prevented well-intentioned policy makers in the Ministry of Education, such as Yagodin – the last Soviet Minister of Education, and Dneprov – the Minister of Education of the new Russian Federation – to implement their agendas (Reilly, 1996; Webber, 2000).

Perhaps with more time to realize their strategies, the reformers in the education policy arena would have been able to institutionalize and bring to scale the pockets of innovation present in the country, despite the dwindling financial and material resources. However, the political changes that ensued, as well as the deepening economic crisis and failure of the state, shifted education down the agenda of policy makers. Furthermore, the education system of the Soviet Union fell apart along with the empire, leaving it up to the local elites and *intelligentsia* to carry out reforms while at the same time coping with ongoing problems.

2.1.2. THE 1990'S: THE DEPTH OF THE CRISIS

It would not be an exaggeration to state that the profound political, economic, and social crisis that followed the breakup of the Soviet Union was one of the worst in history, lasting as long as a decade and leaving deep scars in the memories of millions of citizens who endured through the hardships of the 1990's. For some countries, the end of the Soviet period meant the beginning of democracy building, while for others it marked the beginning of a new form of authoritarianism and control. Still others, like Russia or Kyrgyzstan, started out as fledgling democracies, only to find the newly established democratic

freedoms gradually fading away, as their autocratic leadership tightened its grip on power. While in the early 1990's all of the post-Soviet states experienced severe deficits of cash resulting in delays in teacher and staff salary payments often months on end, as well as a lack of financing to cover basic maintenance needs (DeYoung and Suzhikova, 1996), by the end of the decade the widening variance in economic performance across the post-Soviet region became apparent. Countries with strong export-oriented economies, such as the Baltics, or substantial natural reserves, such as Russia, Azerbaijan and Kazakhstan, fared better than some of their neighbors: with growing national budgets, governments were able to raise the salaries of public school teachers (who were mostly civil servants), carry out large school infrastructure projects, and embark on new initiatives supporting the improvement of quality in public school systems (MOE Russian Federation, 2009; MOE Kazakhstan, 2009).

Despite the differences in economic capacity, however, reforms in the education sector took place across the former Soviet space. Curriculum was the natural first target for reform, with Soviet-era content replaced by the new narratives of the newly independent states as part of the de-ideologization efforts (Webber, 2000), if only for subsequent re-ideologization with values reflecting the newfound national identities. As an illustration, Kissane (2005) describes how the history curriculum underwent de-Russification and de-Sovietization in post-Soviet Kazakhstan in order to fit the goals of the new national ideology, with an emphasis on nationalizing the political discourse and adopting a critical perspective on "colonization" by the Russian Empire. Similar efforts took place in Russia proper, introducing "new images" of Russia, including a favorable image of the late Russian monarchy and a critical discussion of "what went wrong" in the Soviet period (Shevyrev, 2005). Other nationalizing efforts in the newly independent states included language

policies put in place for compulsory education, where the stiffness of national language requirements ranged depending on the stance of the national elite towards Russia (Laitin, 1998). In some cases, however, the previously stringent and uncompromising policies adopted by the new governments towards Russian language education was subsequently reversed in line with the "language of the new allies" (Silova, 2004).

Reforms also touched other aspects of the education systems in the newly independent states, such as financing, management, and quality control. Previously entirely State-supported, schooling systems began to incorporate private education models, at all levels of education - from preschool to university, although the majority of private educational entities appeared at the tertiary level. Furthermore, fiscal pressures on the national governments along with the emphasis on neoliberal policymaking advocated by international donors served as an impetus for decentralization of financial and management burden towards the local level. In Central Asia, for example, the decentralization of fiscal burden without a corresponding increase in local government discretion in management decision making left many municipalities without the resources to adequately maintain their schools, and led to widening disparities in the quality of education between the wealthier urban communities and the poor rural ones (Mertaugh, 2004). Nonetheless, as Chapman et al. (2005) note, quality of learning was a national top priority for all of the Central Asian countries, with substantial effort allocated towards the development of new standards, revisions of curricula, and the development and publication of textbooks. Teacher education and monitoring of teaching quality were generally recognized, but seldom resulted in actual plans of action or specific efforts. Quality assessment in general was focused mainly on measurement of learning outcomes, with a specific emphasis on the development of standardized tests, which served a dual purpose, acting as a guard against

corruption of high-stakes testing and examination. However, investment in teacher capacity remained a priority in rhetoric, but not in action (Ibid.).

An area that has been generally left untouched by reform, despite its profound effects on the quality of teaching and the potential for innovation to take root is the teacher salary structure. The system created during the Soviet period was based on the principle of classroom hours load – *stavka* – which reflected the number of hours that the teacher was expected to spend in front of the class. All of the other responsibilities of teaching, such as checking homework were either paid separately, or not at all. The number of teaching loads that a teacher could take was not limited to one, resulting in an incentive for teachers to increase hours to maximize take-home pay, often at the detriment to their ability to prepare for class or engage in professional development. A recent UNICEF study (forthcoming) problematizes the situation the following way:

> *The low teacher salary in combination with the low statutory teaching load (18-22 hours) have turned the teaching profession into a part-time job, encouraging teachers to either take on additional hours in the school, seek for additional income from parents, or take on additional job outside the school (p. 11).*

Such conditions, where teachers are underpaid and are incentivized to spend more hours in front of the class, make it difficult to expect substantial interest in innovative teaching methodologies, which by nature require a greater investment of time and resources than traditional direct instruction. This topic is explored in Chapter 7, where interviews with teachers in Kazakhstan reveal the extent of this challenge.

A number of reforms took place at the tertiary level, as governments sought to eliminate corruption by introducing standardized admission procedures to public universities, and took steps to bring the structure of higher education closer in line with perceived exemplars from the West (Drummond and DeYoung 2004; Pliksnys et. al. 2009;

Makarova and Solomennikov, 2008). Structural changes have not been as profound, however, with only a handful of countries addressing the financing mechanisms, such as Russia or the Central Asian states of Kyrgyzstan and Tajikistan, and few incremental adjustments made to teacher payroll and workload norms. Government-led reforms of teacher training and instructional practices have been even less noticeable, despite the official recognition of the need to change pedagogy in order to meet the demands of the modern labor market and civil society. International educational nonprofits, however, established a presence in the region in the early days of the post-Soviet era, and have been actively promoting their models of student-centered instruction directly to the schools, offering professional development for teachers and small grants for teaching aids and supplies. Programs rolled out by the Open Society Institute of the Soros Foundation for instruction at the elementary and secondary grades, in particular, reached out to broad networks of teachers, who were then organized into formal professional membership associations to guarantee sustainability beyond the initial Soros Foundation funding (Silova, 2008). Large donors, including USAID and World Bank, also emphasized in-service training in their packages of technical assistance for education. Large-scale, multi-country programs such as, for example, PEAKS – funded by USAID and implemented in Central Asia by a consortium of international education nonprofits, had in-service training in student-centered methodologies at the core of its comprehensive strategy in each of the recipient countries.

Perhaps it is due to these initiatives of non-state actors, such as international NGO's and donor agencies active in instructional reform, that the new national governments have largely incorporated an emphasis on student-centered pedagogy and development of critical thinking skills in the national strategies for education. In Russia, the new 2004

curriculum emphasizes the development of "key competencies, personal creative development and... application of knowledge and skills" (Kovaleva and Krasnianskaya, 2008). Russia's federal Ministry of Education also disseminates competitive grants to schools and teachers who demonstrated innovation in teaching (Ministry of Education of the Russian Federation, 2006). In Armenia, the development of "basic competencies, cognitive and creative potential, and independent problem-solving skills necessary for decision making in the market economy" is similarly a key goal of the 2004 national education curriculum (Mkrtchian, 2007). Changes also took place in Lithuania: according to Bigeliene, Elijio, and Strickiene (2008), in the country's math and science instruction "the academic way of presenting information was replaced by a more student-oriented teaching approach that took into account the age and experience of students" (p. 366). Similar principles and concepts can be found throughout the national education strategies of Georgia, Ukraine, and other countries in the former Soviet Union. In Kazakhstan, one of the goals of the 2005-2010 national education strategy is to change "the role of the student in the process of learning… from passive recipient of information to active participant of the education[al] process" (Government of Kazakhstan, 2004). Thus, student-centered instruction became the core element of change and transformation in the education policy talk throughout the post-Soviet space. Kazakhstan, as a state with a very strong executive power, makes for an interesting case of observing whether such rhetoric reflects genuine political will, and more importantly, capacity for reform of instruction.

2.2. KAZAKHSTAN: CHANGING THE NARRATIVE, CHANGING THE CLASSROOM?

In 2007, Kazakhstan for the first time participated in an international assessment of student achievement, and showed remarkably high performance of 4th grade students in

mathematics and science, with its average score ranking in the top ten on both subjects. These outcomes became a source of great pride for the government of Kazakhstan, and prompted curiosity in the potential relationship between the state education policy and implementation, and student achievement – a usual dynamic applied to the understanding of high-performers on international student assessment studies. In this dissertation, I will not explore the causes of these academic successes, but in Chapter 7, I will explore the current context in which instructional changes take place, and offer insights that may help understand the results of the quantitative analysis of its TIMSS data. In this chapter, I provide an overview of the latest developments in Kazakhstan, as well as the connection between this case and the theoretical framework outlined in Chapter 4.

2.2.1. OVERVIEW

Education policy in the Republic of Kazakhstan is defined by National Education Development Programs (NEDP), which are strategic documents that outline the challenges and opportunities facing the education system for a period of several years, and set policy goals to be achieved during that period. The most recent National Program covered the period from 2005-2010; the newly adopted program reaches over a longer time span, from 2011 to 2020, dividing the ten-year period into two phases, five years each. The 2005-2010 program set ambitious goals for a complete transformation of the education system, from an expansion of school construction and increases in teacher status, to "changing the mindset of the teacher", advocating a shift towards student-centered, competency based learning. A major structural change - the addition of the 12th grade, increasing the duration of upper secondary school - was also planned to be implemented during this time period. However, as the 2009 Annual Report on the State of Education in the Republic of Kazakhstan (MOE 2010) acknowledged, many of the challenges remained at the end of the

strategy period, and had to be postponed by several years due to inadequate capacity and lack of resources to implement stated policies. Nonetheless, the substantive improvement in the availability of resources in the public education system during the past several years is indisputable. At the same time, this period saw an expansion of accountability mechanisms, where teachers, schools, districts and regions all followed a hierarchical system of assessment and control. The competitiveness of the country's labor force was set as the ultimate goal of the education system by President Nazarbayev, and schools are bearing the lion's share of the burden of bringing up a capable and versatile generation.

The system of teacher training, both pre-service and in-service, have largely remained unchanged from the pre-independence era. Pre-service teacher training consists of two tracks: a two-year program at a special teacher training college (called *uchilische*), which admits students after completing the basic compulsory education of 9 grades, and provides a certificate for teaching primary grades; and a four-year program at four-year university, which admits students after 11 grades of secondary school, and provides a certificate for teaching at the secondary level. Teachers trained at four-year program also receive a general Bachelor's degree, while specializing in one or two subject matter subjects (i.e. Kazakh language, Math, Physics, Chemistry and Biology, etc.). In-service training is required for continued certification every five years and it is offered by state-in-service training institutes, with courses both in refreshing content knowledge, as well as in new requirements for curriculum coverage, changes in textbook requirements, and other updates.

Quality control was formally institutionalized in 2004 through the creation of the National Center for the Assessment of Quality in Education, a sub-agency within the authority of the Ministry of Education. Like most agencies in Kazakhstan, the Center is

structured hierarchically with a national body in the capital, and local branches in major cities and regional centers, housed either in the local in-service training institutes or offices of other branches of the Ministry of Education. However, while the Center for Quality in Education is responsible for a substantial amount of data collection for monitoring purposes, it is not the monitoring information that carries the biggest stakes for schools in Kazakhstan. During the past seven years, the country has expanded the mandate of the National Center for Testing, mainly through the universal administration of the Unified National Test (ENT) – a two-in-one graduation and university admission exam. In 2010, all students intending to apply for admission into any of the domestic institutions of tertiary education were required to take the exam.

Math and science education, and particularly, gifted programs in math and science, is an area of particular interest of the government, and of substantial budgetary support. Gifted education in general, either in the form of tracks within regular comprehensive schools, or as separate schools for gifted students, has expanded in recent years. The Ministry of Education reported a threefold increase in the number of students enrolled in specialized and gifted programs between 2004 and 2008, while the number of gifted tracks within regular schools went down –a trend that signals institutionalization of tracking in general secondary schooling. In addition, a new network of gifted high schools with direct government support - the Nazarbayev Intellectual Schools - was launched in 2008, with an eventual target of twenty schools to open by 2020, according to the new National Education Development Program. These schools will serve as the cradle of the future elite of the country, as well as a major building block of its competitiveness on the world labor market.

The 2011-2020 National Education Development Program calls for a complete restructuring of the education system. With the introduction of the 12-grade schooling

policy, carried over from the previous strategic period, the general secondary school will consist of ten grades with a final comprehensive exam upon completion. Results of the comprehensive exam, taking the form of standardized exams, will be used to determine whether the student continues on to one of the academic tracks in high school, or is requested to continue his/her education in a separate vocational school. According to the MOE, this plan will be fully operational in 2015. The National Program 2011-2020 is rooted in the 2008 proposal from President Nazarbayev, the "Intellectual Nation – 2020", which he first voiced in an address to the recipients of the prestigious Bolashak government fellowship for study abroad. The proposal called on educators to put a concerted effort to raising the quality of learning in the nation's schools, creating greater space for innovation, and bridging the digital divide.

Textbooks in mathematics and science continued to form the core of the curricular content in Kazakhstan, making the choice of the textbook a crucial element of instructional quality. In the schools visited for this study, teachers preferred textbooks from neighboring Russia, even though they were not approved by the State. Russian textbooks were considered stronger in many science subjects in secondary grades. Even some Kazakh language teachers in urban schools used the Russian textbooks, translating the problems into Kazakh for class work. The structure of the textbook also heavily influences the structure of the lesson itself.

2.2.2. STUDENT-CENTERED INSTRUCTION AND TEACHER QUALITY

In his 2008 statement proposing the Intellectual Nation 2020 project, Kazakhstan's president Nursultan Nazarbayev emphasized the importance of educating students to develop "an ability not only to consume, but to create new knowledge and innovation"

(*Kazakhstan Today;* 30.01.2008). He further argued that "the most valuable piece of learning is the ability to think creatively, to transform knowledge, to generate new solutions, new technologies, and [other] innovations". Therefore, said the President, to be able to bring up a generation of such citizenry, the nation's schools have to adopt new methods, new approaches to teaching, recruit new professionals (Ibid.). The other two of the three policy directions of Intellectual Nation 2020 were: (1) the integration of information technologies in education and (2) moral and spiritual education, with the goal of counteracting what was seen as negative effects of globalization.

As can be expected in a country with as strong an executive branch as in Kazakhstan, Intellectual Nation 2020 was adopted and fed into the new national education policy, the National Education Development Program (NEDP) 2011-2020. This new policy document contends that the main challenge of the new era in the country's educational development is to reorient teaching so that it develops new "key competencies", such as critical thinking, creativity, and versatility in the workplace, shifting from traditional consumption of knowledge from the teacher. It must be noted that idea of "changing the mindset of the teacher, and… shifting the role of the student from a passive recipient to an active participant of the learning process" had been voiced in the Government's previous National Education Development Program document, covering the period of 2005-2010 (MOE 2004).

However, neither the 2005-2010, nor the draft 2011-2020 NEDP contains concrete strategies to implement instructional reform, other than pledging to increase teacher salaries and expand access to in-service training opportunities. The provisions of the 2011-2020 document seeks to "raise the status of the teacher" by raising salaries, restricting admission into teacher colleges, reestablishing additional teacher certification procedures,

and introducing teacher choice of the providers of in-service training, through a pre-paid voucher mechanism – which effectively creates a blend of state and market providers for teacher training. Teacher working conditions are to be improved as teachers of some disciplines, such as math, the natural sciences, and languages, receive new equipment and supplies to outfit their classrooms, including lab equipment and new technology in the form of 'interactive boards" – classroom board-size screens connected to the teacher's desktop, with pre-loaded educational resources and access to the internet. In Chapter 7, I will further explore these factors in the policy environment, as I go deeper into the context of the implementation of student-centered instruction in Kazakhstan.

2.3. CONCLUSION

In this chapter, I offered a brief overview of the background and context of the education systems in the post-Soviet region, and provided greater depth of detail on the most recent developments in Kazakhstan. I showed that the Soviet education system had its strengths and its weaknesses when it came to instructional methodology, and that there were innovative educators and practices of pedagogical excellence even during the height of the Soviet period. However, they never went mainstream, and the education system remained largely untouched until the late 1980's and early 1990's. I also discussed the political and economic conditions of the post-Soviet crisis that affected the education systems in Eurasia, and noted the stark differences between the former Soviet republics nearly twenty years after they gained independence.

I also briefly described the political and policy environment in Kazakhstan, and the prominent place allotted to student-centered instruction, and "placing the student in the center of the learning process" as key to the change and transformation of the education

system. The rhetoric closely resembles that of other state-driven reforms, as will be described in Chapter 3. Finally, I showed why drawing parallels between this post-Soviet context and the case of Cuba is appropriate, and why the theoretical framework drawn on the basis of the Cuban case was applied in this dissertation.

CHAPTER 3.
LITERATURE REVIEW:
THE CONCEPTUAL FOUNDATIONS

3.1. INTRODUCTION

In this chapter, I offer a brief review of literature on the two main topics interwoven in this dissertation. First, because my dissertation examines the relationship between student-centered instruction and achievement outcomes, I begin with a look at the key historical contributions to the theory of learning behind student-centered instruction, as well as some practical applications of this theory. I also review some examples of evaluating the impact of student-centered teaching in comparison with traditional methods of instruction, with a few small-scale studies based in the United States, and some broad, large-sample studies done using data from international assessments. Noting the ambiguity of some of the findings, and the lack of a conclusive result across different studies, I offer potential explanations of ambiguity.

Secondly, I turn to a few cases of successful education systems, some of which are highly centralized, and others allow for substantial freedom to the schools, to help set the context for understanding the role of the state in achieving academic success. I find commonalities across the cases, and give particular attention to the theoretical argument for the role of the state embedded in one comparative study, in which Cuba is juxtaposed against other education systems in Latin America. I conclude by summarizing key insights from the chapter, which serve to set the main parameters for my theoretical framework and the methodology for this dissertation.

The review offered in this chapter is by no means exhaustive, but is intended to help the reader understand the foundation behind the main elements of this dissertation: one,

student-centered instruction and its expected contribution to learning, and two, the importance of the environment in which it is implemented.

3.2. STUDENT-CENTERED INSTRUCTION

Student-centered instruction is a term almost ubiquitous in global debates in education, denoting a fundamental shift from a traditional "teacher-centered" dynamic to a more democratic and open environment in the classroom. This change in pedagogical approach is not without implicit judgment, and student-centered instruction is seen as more "modern" and likely to instill civic and democratic values, as well as stimulate progress, while the persistence of traditional methods are associated with historical backwardness. In the post-Soviet region, international donors and NGO's working education focused on changing instruction, training teachers to use methods and techniques intended to help students develop critical thinking skills – such as discussion, reflection, project-based learning, use of graphic organizers, experimentation, and other methods. Projects such as PEAKS (USAID, 2003-2007) funded by the U.S. Agency for International Development, for example, left an institutional legacy in Central Asia by forming a group of professional teacher trainers with capacity in student-centered methods (Steiner-Khamsi et al., 2007). Another major actor in post-Soviet education, the Open Society Institute, has made it its mandate to change instructional practices in elementary and early secondary grades, by training teachers in the use of packaged methodologies such as Step by Step and Reading and Writing for Critical Thinking, designed to stimulate the development of critical thinking skills and the dissemination of democratic principles of social organization (Silova, 2008).

However, no claims were made by these international actors about the impact of this shift in pedagogy on student achievement – the emphasis was on non-cognitive aspects

of schooling. Nonetheless, the learning theory behind student-centered instruction does argue that better student outcomes can be expected if the emphasis of teaching is on the student, and greater social interaction is introduced in the learning process. Results, however, are mixed, making it difficult to argue that the theory has proven itself in this aspect of learning. In this part of the chapter, I provide an overview of this literature, followed by a series of cases that show how state education systems have adopted this shift at the level of national education policy.

3.2.1. KEY HISTORICAL CONTRIBUTIONS

The theory and practice of student-centered instruction is by no means new. In the more recent history, a profound influence on pedagogy can be traced back to the writings of Jean-Jacques Rousseau, who put forth a fundamentally different vision of education, based on the belief that schooling should nurture the child's innate motivation for learning, rather than instill discipline and order (Bowen, 1980; Falk, 2006). While it is the Progressive education movement of the early 20th century and the subsequent developments in constructivist theory in educational psychology that are most often credited as the foundation for the contemporary "varieties" of student-centered and active learning instructional methods, the roots of student-centered instruction were also present in the Russian and Soviet pedagogy of the 20th century. The literature on this topic is immense, and by no means would be adequately represented in this chapter. I therefore focus on the major contributions to theory of student-centered learning, made by scholars that left a lasting legacy in the field.

In Western educational tradition, John Dewey, one of the most influential scholars and philosophers of his time, argued that education is a means and an end in itself, a process

of continuous learning and self-perfection of individuals for the better of society as a whole (1916). In Dewey's theory, education and social life were inseparable; and in order for education to be meaningful to the child, it must be situated in connection with the child's social and home life, and schooling should help children make meaning of their role in the home and society. Education must use the child's capacities and interests as a starting point and continuous reference framework, as a way of establishing continuity of learning. The teacher is not there to impose ideas or form habits, but to help form the overall learning environment for the child, and to assist the child in engaging with outside influences and responding to them. Education is therefore by definition child-centered, focused on the creative and learning potential of the child, and rejecting uniformity in any aspect, be it curriculum, teaching, or grading.

The theory of constructivism – another layer of foundation for student-centered education –builds on the works of the Swiss psychologists Jean Piaget (Montagnero and Maurice-Nauville, 1997). Piaget theorized that intelligence develops through the construction of relationships and linkages between oneself and the surrounding environment. Furthermore, development of intelligence progresses through a sequence of chronological stages corresponding to the age of the child (sensorimotor, preoperational, concrete operational, and formal operational) characterized by increasingly complex ability to construct knowledge. Piaget's contribution was mainly to the field of psychology, and yet it had profound implications for educational practice. Since intelligence develops through the construction of mental structures between the subject and the object, education, according to Piaget's theory, is most effective when it stimulates the development of natural intelligence by creating an atmosphere conducive to experimentation and an active role of the child in constructing his or her reality and learning, through direct interaction with the

environment. A stimulating educational environment allows substantial place for the child to engage in sensorimotor activity, manipulation of objects, and structured play, with positive encouragement from the teacher (Isaacs, 1972; Iganaki, 1992). Educational models based on Piaget's theory are built as open spaces where the teacher does not provide the correct answers, but stimulates children to engage and manipulate objects, make predictions about the outcomes of experiments, and compare and analyze their predictions with actual observations (Isaacs, 1972).

When it comes to practical implementation of child-centered pedagogy in Western education, the system developed by Maria Montessori was perhaps the most revered and widely replicated (Standing, 1962). Manipulation of objects, mobility in open classroom space, and freedom in deciding the learning activity for the day formed the general framework of the Montessori principles. Education in the Montessori method is built on the child's own interests, and the classroom materials and teacher presentations are designed to stimulate interest in new topics and activities (Lillard, 2005). The Montessori method rejects extrinsic rewards, such as prize and punishment, as unnatural and detrimental to learning, and puts a premium on intrinsic motivation - the natural interest in a subject or topic - to ensure that children accomplished their goals and behaved well in the classroom. Evaluation is entirely absent from this method. Lillard (2005) finds evidence in modern research in support of the Montessori method, and in particular, cites several studies that found extrinsic rewards as negatively affecting student performance.

3.2.2. STUDENT-CENTERED INSTRUCTION IN EURASIA

Scholars and educators in Russia and the Soviet Union also made important contributions to student-centered instruction as a strand in pedagogy in early and mid-20th

century. In fact, one of the major theoretical foundations of active learning was developed by the Russian psychologist Lev Vygotsky in the 1920's. Vygotsky's theory on the zone of proximal development (Vygotsky, 1978) emphasizes the distance between what a child can do on their own, and what they can do with assistance from an adult or "more knowledgeable others" and in a supportive environment. This area marks the child's immediate cognitive potential, and therefore, the construction of such an environment is crucial for the child's cognitive growth. Vygotsky saw cognitive development as a progression from lower mental functions, such as perceptions and associations, to higher-order mental functions, such as language, voluntary attention, and problem solving. This progress takes place through social interactions, where children first experience more complex notions and forms, and eventually internalize their experiences into their own mental functioning. Complex tasks that lie at the upper bound of the zone of proximal development, are first demonstrated to the child, who then strives to accomplish them with assistance from others, and eventually is able to complete them completely on his or her own without assistance. In Vygotsky's words, "what lies in the zone of proximal development in one stage is realized and moves to the level of actual development in the second" (Vygotsky, 1987, p. 11., as cited in Doolittle, 1995).

Vygotsky's ideas regarding cognitive development as a social process, in which "more knowledgeable others" play a crucial role in the advancement of the child through increasingly complex mental functions, place an emphasis on forming collective learning forms. However, it was Anton Makarenko (1888-1939) that most often associated with the idea of education as part of a collective effort. Makarenko made a profound influence on Soviet education, as he emphasized the transformational potential of work and group conscience through his work with delinquent youth. However, while Makarenko did place a

substantial emphasis on a work ethic and education as pathway for building a better society by overcoming challenging circumstances (Holtz, 2002), his methods could hardly be characterized as student-centered: his insistence on strict discipline and intense group pressure are contrary to the freedom and individualism that lies at the core of student-centered instruction.

Another influential educator in the Soviet Union was Vassiliy Sukhomlinsky (1918-1970), a writer, teacher and school principal of thirty five years. An education practitioner, Sukhomlinsky created a model school that reflected his principles of pedagogy. Sukhomlinsky's school placed the child in the center of the learning process, and his/her interest and self-realization at the heart of pedagogical practice. He stressed the importance of education to adapt to the needs and specific interests of the child, such as not to inhibit but to help the natural process of intellectual development. Sukhomlinsky adamantly rejected the concept of punishment, which he considered as not only oppression and offense on the dignity of the child, but also a severe assault on the child's future desire to improve morally and intellectually (1980, as cited in Papadopoulou, 2008). Sukhomlinsky believed that the educator should meet the student at the level of the child, and establish personal trust, which would then bring about a genuine transfer of knowledge. These ideas are at the core of student-centered instruction, both in Russia and the former Soviet Union, as well as the West.

While Sukhomlinsky's principles were reflected in his concrete model of a school, they never quite reached the mainstream in Soviet education. A number of educators, however, pushed for a reform of instruction, and called for a more student-centered approach to teaching. By the 1980's, innovative educators shared their experiences in mainstream teacher's newspapers and journals, publicizing evidence of success (Webber,

2000); however, the subsequent political changes and prolonged economic crisis prevented scale-up of innovation.

Despite the wide differences in the starting points for the theories discussed above, as well as the Western theories of child-centered education, they share substantial common ground. These theories shift the focus of education from the teacher and the school to the student – the student's needs, capacities, interests, and the formation of complex relationship between the child and the world around him. They all argued that a supportive environment and a teacher willing to assist the child in the making of knowledge is crucial for making a truly meaningful educational process.

However, all of these pedagogies were devised in small contexts, in highly individualized environments, and with well-trained and intrinsically motivated teachers. It is a different question, then, whether student-centered instruction works equally well when it enters mainstream schools, with its high variability in the quality of teachers, the level of resources available at schools, as well as the value placed on education by families and their ability to support their children. Furthermore, even if student-centered learning is beneficial, it is possible that these benefits could be entirely non-cognitive – or, to be more precise, not measurable in terms of academic achievement. In order to formulate the methodology for the study, I examine how student-centered methodology is operationalized in mainstream schools, and look at available evidence of the impact of such methods on student academic achievement.

3.2.3. THE PRACTICE OF STUDENT-CENTERED INSTRUCTION: WHAT IS THE EVIDENCE?

Since the theoretical underpinnings underlying student-centered instructional methods argue that placing the child in the center of the educational process will result in better learning, the validation of these theories calls for a comparative analysis of achievement outcomes in traditional and child-centered settings. After all, instructional approach is but one slice of the potentially large "pie" of factors that influence teacher effectiveness in the classroom. Teacher content knowledge, for example, has been the subject of extensive research and theorizing, both in the United States and in comparative perspectives, and not surprisingly, in-depth knowledge of subject matter in mathematics and science is paramount to teacher success in the classroom (see, for example, Ball, Lewis and Thames 2008; Schmidt et al., 2007). Furthermore, a strong and rigorous set of curricular standards were also identified and extensively documented as strongly related to higher student performance specifically on TIMSS math and science assessments (Schmidt et al., 1997). However, neither of these arguments rules out the potential effect of instructional methods.

The instructional models built on the theoretical foundations for student-centered instruction, some of which are discussed above, involve continuous reflection and adjustment of approaches depending on the developmental needs of the child, in order to maximize the realization of his/her potential for cognitive and creative development. Early students of such models juxtaposed traditional, direct instruction with their creative environment, and, based on this juxtaposition, argued for the superiority of student-centered methods. Gradually, the literature has seen the growth of more rigorous research to evaluate the impact of student-centered instruction, predominantly in the form of case studies with direct observation of teachers and students in different settings.

Naturally, there is hardly one set of techniques that are consistently used by all teachers to form a recipe for student-centered instruction. As one example, a set of practices generally used by *effective* teachers of mathematics and science was compiled by the National Research Council in their volume "*How people learn: Brain, mind, experience, and school*" (Bransford et al., 2004). Since the authors emphasize student-centered instruction as the most beneficial methodology that follows research on cognitive development, the techniques highlighted in this volume as "effective" appear to be one of the clearest articulations of student-centered teaching available in the literature. The activities included in the volume are as follows (p.169-189):

- Students spend a great deal of time discussing alternative strategies with each other, in groups, and as a whole class.
- The teacher challenges students to think and make sense of what they are doing in math, invent their own strategies for solving problems and discuss why those strategies work.
- Students are prompted to think about aspects of everyday life that are potentially relevant for further learning.
- Interactive science demonstrations are used to overcome the students' erroneous beliefs about scientific phenomena.
- In science, students learn to develop a line of argumentation that uses what one has learned to solve problems and explain phenomena and observations.

Darling-Hammond et al. (2008) provide a broad review of research in effective teaching, based on case studies conducted in the United States. They find evidence for the positive effect of the principles of student-centered instruction. For example, they find that students learn more deeply when they can apply classroom-gathered knowledge to real-world problems, and when they to take part in projects that require sustained engagement and collaboration. They also argue that active learning practices have a more significant impact on student performance than any other variable, including student background and

prior achievement - making it possible for schools to overcome the baggage of SES. The authors find evidence among case studies that students are most successful when they are taught *how* to learn as well as *what* to learn. In line with Piagetian principles, the authors, citing research by Newmann (1996) argue that a positive impact on learning results from students participating in lessons that require them to construct and organize knowledge, consider alternatives, engage in detailed research, inquiry, writing, and analysis, and to communicate effectively to audiences.

Small group instruction, or cooperative learning, specifically, has been a subject of numerous studies. However, most of such studies are performed by advocates of student-centered instruction, and are conducted in settings where randomization of the methodology or approach in question is difficult to ascertain. For example, Johnson, Johnson, and Stanne (2000) carried out a meta-analysis that examines the effects of cooperative learning. They identified ten stand-alone programs in cooperative learning that have been designed and disseminated in the United States since 1960's to 1980's, and estimated separate effects for each of the programs. The authors found that most of the cooperative learning strategies had statistically significant positive effects on student learning, in some cases as large as 0.83 standard deviations, when compared to competitive methods, and 1.03 standard deviations when compared to individualistic method, although not all cooperative learning strategies were equally effective. Such large effect sizes are quite unusual in education, and one would be correct to use caution in interpreting the results of this study: it is not clear in the meta-analysis whether the selection bias was appropriately minimized, and whether other potential explanations (such as teacher competence and/or motivation) could be convincingly ruled out.

While some positive evidence was registered in small-sample studies in the United States, larger studies using unified measures of instructional methods and achievement outcomes, such as international achievement studies, have yet to show a positive effect. Consistent positive relationships between active learning instruction and achievement outcomes have not been found. For example, in its study of top performers in PISA 2006, a large-scale international assessment of 15-year-olds, OECD reported that the top performing students reported *lower* levels of engagement in hands-on, investigation, and student interaction activities in their science lessons than did the lower-performing students (OECD, 2009). In looking at the effects of separate instructional methods in math and science on the student TIMSS performance in these domains in three East Asian countries, two studies found both positive and negative effects associated with student-centered teaching methods, with the more participatory methods producing negative or insignificant coefficients (House, 2005; Leung, 2002, as cited in Aypay, Erdogan, and Sozer, 2007). In Turkey, using results from TIMSS 1999, another group of researchers found statistically significant negative relationship between student-centered instruction in science classroom and science achievement: the lower performing schools were also ones where teachers were more likely to use student-centered pedagogy in their lessons (Aypay, Erdogan, and Sozer, 2007).

In sum, the evidence of the impact of student-centered instruction to date has been mixed, ranging from very large and highly significant results from meta-analyses of mainly U.S.-based research, to barely significant or negative in studies involving large international achievement databases. What can explain such contradictory evidence? There are several plausible explanations for this discrepancy. First, the methodologies for both studies carry their own flaws. In the case of meta-analyses, studies of varying quality, different

dependent variables, and validity are mixed in one bag, making it very difficult to discern reliable information about the relationship in question. In the case of quantitative studies mentioned above, the principal methods used for the analyses (OLS regression and discriminant factor analysis) do not allow for a correction of bias associated with nonrandom assignment of students into classrooms that do or do not experience student-centered teaching methods (Gelman and Hill, 2007). In this situation, it is impossible to separate the effect of the student background and school variables from the effects of the instructional methods themselves. In other words, no potential alternative outcomes resulting from a different treatment regime could be estimated for students. Another potential explanation may be that the measures of student achievement are measuring completely different things, and not one and the same construct. Further, assuming the methodological weaknesses away, true effects could be different in different settings - in this case, in different countries: positive in the U.S. and negative elsewhere. This explanation would imply that there is something fundamentally different about the students or teachers between the participating countries (but not different within the countries) that strongly influences the effects.

Finally, one should not discount the importance of system-level factors, such as the role and importance played by the central state, the level of cohesion or coupling between elements of the system, and the presence of a shared frame of reference among teachers and administrators about what it is that education is expected to produce, and what measures of success are used to judge the effectiveness of their work. The knowledge and competence developed by the pre-service training system in a given state is also a factor of crucial importance, however it is often left unmeasured, particularly in large scale studies. Some education systems have been consistently successful on multiple international

assessments, and understandably generated interest on the part of the international education community in their policies, their enabling conditions, and the construction of the relationship between the state and the individual – all in the hope of replicating their success in a different environment. In the rest of this chapter, I briefly examine some of these successful state systems, in order to then proceed, in Chapter 4, to building a theoretical framework for understanding the implementation of student-centered instruction in Kazakhstan.

3.3. WHAT ABOUT CONTEXT? ROLE OF THE STATE IN SUCCESSFUL SYSTEMS.

As I mentioned above, in this dissertation I intend to not only examine the statistical evidence of impact of student-centered instruction on achievement, but also explore the context of the education system in the post-Soviet nation of Kazakhstan. In particular, I am interested in whether the structural characteristics of this highly centralized education system have played a role in how student-centered instructional approaches have been adopted and disseminated across schools. Alongside this main interest, I also look for indications that the state education policies and structural characteristics had affected the performance of Kazakhstan's fourth graders in math and science, as it was measured in 2007 by TIMSS. The level of concentration of power and authority in the hands of the central government is perhaps one of the most visible defining features of Kazakhstan's education system, and it stems from the broader political and social landscape in this Central Asian country. In this part of the literature review, I offer a brief look at other highly effective and high-performing states (in terms of their mean achievement on international tests), with the intent of setting a frame of reference for a theoretical framework that I will later use in understanding the results of my quantitative and qualitative data analysis in

Kazakhstan. The countries included in this review are: Singapore, Hong Kong, Japan, Finland, and Cuba. The choice of these cases was driven by their strong performance on international achievement tests, as well as the presence of a visible, if not always the same, role of the state in their education system. They differ in geography, population size, economy, and a multitude of other factors, but as this section shows, they do possess some similarities in their approach to reforming education, which signifies the presence of a global lingo for education reform, and points to ways that it was adapted to local situations.

3.3.1. SINGAPORE

In their comprehensive overview of the structure and recent developments in the Singaporean education system, Sharpe and Gopinathan (2002) emphasize the central role allotted to education in the country's vision of national economic development. The early goals of the state in the postcolonial period were to create a unified education system that would foster social cohesion in the ethnically and politically divided country, and to build a labor force that would put the country on par with economically developed nations. As the authors put it, "From the beginning, the education system was used as a major vehicle in nation building, with the state acting as `strategic trader' to align provision with the needs of the economy and social cohesion…" (p. 154). Centralization and standardization were the solutions to the problem of division and disparity: through centralizing and standardizing the education system, the state was able to make the education serve its nation-building and industrialization agenda, according to Sharpe and Gopinathan (2002). The authors emphasize the comprehensiveness and the presence of a clear strategy in the ways that the Singaporean government defined the curricular goals, developed instructional materials, developed training and capacity building programs connected to the core curriculum for teachers and administrators at all levels, and introduced regular student

assessment systems during the first three decades of independence. The authors call this stage, starting in the 1950's to the 1980's, the *efficiency* stage in Singaporean educational development, and argue that the remarkable success of the state efforts in building the country's education system was crucial establishing a fundamental level of confidence of the public in its governing elite. "The state established itself as a strategic player", argue the authors, and in turn, gained political legitimacy and long-term stability. The civil service – the core element of the state – established itself as a capable and competent agent of transformation in the eyes of the broader society.

The recession of the mid-1980's, however, brought a wave of neoliberal rhetoric, and saw the beginning of a restructuring process within Singaporean education. The high level of central control in education was now seen as an impediment of the education system, in comparison with Western models such as the United States and the UK. The government began to call for the creation of independent schools, and sought to "decentralize the education system and locate the nexus of reform within the schools themselves" (p.156). School autonomy became an important objective, with the goal of making school-level decision making more relevant to the needs of the local populations. However, the authors still observe that while a process of decentralization, in the form of autonomous schools, did take place in the 1990's, the central Ministry of Education not only retained, but further solidified its role in defining the parameters of change, and designing and leading every reform in the education sector in the country during this period. New and experimental models in instruction, with an emphasis on critical thinking skills and creativity, firmly entered the debate on the new directions for Singapore's education system.

This new approach to instruction was manifested most clearly in the "Thinking Schools, Learning Nation" (TSLN) initiative, launched by the government of Singapore in 1997. The initiative, first voiced by the Prime Minister of Singapore, was later translated in a series of measures designed to make the education system more responsive to the needs of the knowledge economy, to expand the definitions of achievement outcomes. Project-based learning and experimentation were stipulated in the national curriculum, as was the importance of "creative energy" as an outcome of schooling. In addition, considerable resources were expended into the provision of computers and internet access for the schools, and the state envisioned increasing the proportion of instruction with the use of computers to 30% of curricular time (Ibid.). Mok (2003) states that the TSLN vision can be seen as a strategy of the state to prepare the nation for the challenges of globalization.

One of the state-led measures as part of TSLN was the Schools of Excellence Model (SEM), which introduced a system of self-appraisal for the schools that includes broader measures of performance in addition to the students' academic performance. Mok (2002) underscores the change of the assessment dynamic in SEM: a series of external inspections were replaced by an internal process, thereby creating a sense of ownership of performance results on the part of the schools. The SEM framework assesses not only the end results of the learning process, but also the "enablers", or the human capacity and management conditions required for effective instruction to take place in a school (Ibid.). The internal evaluations within SEM are validated every five years by an external commission, which helps schools define targets and identify weaknesses that hinder their progress. In sum, the criteria of quality were broadened, and more decision-making was shifted to the lower levels of the education system. However, both Sharpe and Gopinathan (2002) and Mok (2003) contend that the central state was as present as ever in shaping educational

development in the country – if not through direct implementation from the top to the very bottom, but through extensive criteria for assessment and accountability, which prompted Mok (2003) to dub the process a new "recentralization" of authority. Sharpe and Gopinathan (2002) describe the process the following way:

> *Despite the decentralization measures noted earlier, the late 1990s, can be said to be characterized as the growing realization that the state would need to continue its role as strategic trader if desired changes in knowledge, attitudes and dispositions were to be achieved. It would need to specify much more clearly than before what the education system was trying to achieve and in particular the kinds of talents required by the new knowledge economy; what changes to the assessment system were needed; what the implications were for school organization and assessment; and what the likely consequences of its policies were. The MOE has moved to bring its own thinking and practice in line with new economic and social realities* (p.160).

As this brief description of Singaporean education system shows, the state invested heavily in the development of strong teaching and management capacity during the first three decades. It later recalibrated its involvement in education, allowing for greater flexibility and an emphasis on creativity and experimentation, and yet its grasp of the education system continued to be strong, with each actor following the vision and strategy outlined by the state.

3.3.2. JAPAN

Another common presence in the top tier across international achievement studies – Japan – is another interesting case of educational change within a highly centralized policy environment. Cave (2001) puts it simply: "It is the government that ultimately sets the education policy agenda in Japan."(p. 177). The author describes a seemingly paradoxical situation in the late 1990's in Japan, when the world was fascinated by Japan as a consistently top performing nation in TIMSS and PISA, and yet numerous commentators in Japan pointed out what they saw as fundamental flaws of Japanese education: the rigidity

and intensity of requirements for academic achievement. The stress level created by the system was connected in the minds of the critics with the growing dropout rates of Japanese high school students, as well as bullying and teen suicides (Cave, 2001). Commentators argued that the Japanese system was too focused on inculcating factual knowledge at the expense of developing creativity and individual inquiry among its students, and was therefore unable to produce a labor force capable of competing on the world economic arena. Fujita (2000) notes that Japan was moving "in the opposite direction from Western nations": instead of resting on its laurels, the state was taking bitter criticism of its educational approach, who argued that it was more important to allow for freedom in education, even at the expense of rigorous academic standards.

As a result, the debate shifted to discussions of neoliberal ideas for Japanese education, and namely the emphasis on diversity, choice, and competition. Business-related commentators were keen on introducing features that they saw as advantages of the education systems of other nations in the competitive labor market. Reports originating in the nation's Ad Hoc Council on Education and the chief advisory agency, the Central Council on Education resolutely rejected the old forms of Japanese education, calling instead for an emphasis on "individuality" and greater freedom for the child in the schooling process, including a slimming down of the curriculum, and interactive activities such as learning through play, and more exploratory and experience-driven teaching methods. Furthermore, according to Fujita (2000), the state was challenged by neoconservative and neoliberal critics alike to undertake reform "for its own sake", in a wave of "reform supremacy" (p.47).

The state responded by reducing curricular hours and including subjects that involve interdisciplinary learning. However, in terms of the structural features of the

system, little has changed by the end of the 1990's, according to Cave (2001) in the ways that the education system was run: despite the fascination with choice as a concept, 97% of the students outside the major metropolises (where the private schools were concentrated) had no say in where they would go to school, and what they would be learning. Nonetheless, Fujita (2000) argued that the prevalence of the neoliberal agenda in educational debate served to erode the "foundation of professional control" built by the state and supported by teachers, and slanted the direction of transformation towards market-driven solutions.

By the mid-2000's, however, the pendulum seemingly swung back to the emphasis on academic skills, and on greater rigor in curriculum (Bjork & Tsuneyoshi, 2005). The "reduced intensity" movement of the 1990's and the loosening standards by the central Ministry of Education created a backlash among a large part of the public, who now believed that quality of Japanese education was now being watered down, and that the emphasis on "21st century skills" such as critical thinking should not have been allowed to displace the basics of subject matter retention. The state backed away from education, and allowed greater autonomy and liberty in instructional approaches, with an implicit encouragement for teachers to use student-centered methods to stimulate critical thinking and creativity. However, the teachers were left to figure out on their own what it that shift actually meant in terms of their classroom management practices. As the public turned back to expect greater academic rigor, teachers found themselves caught between two fires. Bjork and Tsuneyoshi (2005) put it this way:

> *Conflicting messages and expectations regarding appropriate classroom practice have created a stressful work environment for many teachers. On the one hand, the government has not abandoned its calls for an emphasis on teaching that develops thinking skills and other 21st- century abilities. On the other hand, teachers feel compelled, because of pressure from the media and policy makers, to emphasize basic*

> *academic skills, often through drills. And the government's appeals for teachers to devote increased attention to students' "individual aptitudes" have further complicated the situation* (p. 623).

In sum, while the "reduction of intensity" reforms in Japan took place under the leadership of the state itself, the space created during this reform movement quickly closed. The state did not provide leadership for teachers in implementing instructional changes and instead relied on individual choices and market-driven ideas to take root, creating a confusing and disorienting environment for teachers and students, which ultimately resulted in a renewed interest in academic standards and basic learning skills. A large number of private schools and after-school private tutoring institutions, catering to parents who wished for their children to gain a competitive edge in university admissions, also served to undermine the support for a reduced intensity curriculum. The state never completely lost control of the education system, however: the national curriculum, examination, and standards remained firmly within the hands of the central government, and therefore, the burden to fulfill the demands for greater rigor fell on the state.

3.3.3. HONG KONG

The developments in the education system of Hong Kong at the turn of the century have been marked mainly by the tension between the city's colonial past and its present and future as an integral part of China (Morris, Kan, and Morris, 2000; Chan & Mok, 2001). Modeled after the British education system, Hong Kong had an elitist structure of schools, with stringent entrance examinations guiding entrance into high quality secondary schools and universities. The role of the state in education was minimal, and schools were allowed a substantial degree of autonomy, in line with the neoconservative "small government, great market" principle (OECD 2010). Private sector participation in primary and secondary

education was encouraged during the early years of Hong Kong's colonial history, with the state shunning the responsibility for providing education to all its citizens (Morris, Kan, & Morris 2000). OECD (2010) observes that until the 1970's, the state effectively limited the number of students who could obtain upper secondary and tertiary education in Hong Kong, out of concern that an oversupply of highly educated young people in the colonial territory might be breeding ground for social unrest. OECD (2010) argues that the artificial limits placed on higher education enrollments was a rational act on behalf of the government to generate a large supply of manpower for its rapid industrialization in the 1970's and 1980's. Morris, Kan and Morris (2000) describe the process of expansion of educational opportunity not as a state-initiated effort, but as a slow reaction of the colonial government to the inevitable changes in the economic, social, and political environment surrounding Hong Kong:

> *Slowly, and primarily in response to various economic and political crises and social demand for schooling, the state began to control, finance and expand the provision of each successive level of school education*. (p. 247)

While expanding educational opportunity, however, the state did not focus on equalizing the quality of education offered to the newly entering cohorts at the same time. As Chan & Mok (2001) observe, the government of Hong Kong was unwilling to engage in the regulation of quality, and public schools reportedly could differ in the quality of instruction a great deal, even as they had the same level of public financing. Elite schools continued to produce graduates with exceptional level of excellence both in academic subjects and in the arts (OECD, 2010). The pressure on the state to ensure quality education intensified during the decade preceding the handover of Hong Kong to the People's Republic of China: general public anxiety over the transition translated into demands to strengthen and equalize access to quality instruction.

The late 1980's - early 1990's period, quite unsurprisingly, was also marked by an increased emphasis on neoliberal values, and business leaders were increasingly thought to be the main beneficiaries of the education system, and the solution to education problems was thought to be the injection of business models of management into school administration (Chan & Mok, 2001). However, the state did not withdraw from education: rather, private sector management models were brought into the public sector, in a bid to improve the efficiency and cost-effectiveness of the education system.

The period of the state's more extensive engagement with the education sector was marked by the elimination of primary school exit examinations and the development of a new education policy reform document in 2002, which called for a new approach to education, putting the focus on the learning process per se, rather than just on the end outcomes of schooling. This new policy, according to OECD (2010), was informed by "contemporary theories of learning" (p. 152), and in essence, required a shift to student-centered teaching practices and a greater emphasis on the construction of knowledge and the development of critical thinking skills, as well as an explicit move away from memorization.

The proclamation of the new approach to education involved the development and roll out of the new curriculum for primary and secondary grades, introduced gradually through a controlled process of "perception management" managed by the government. Every school in Hong Kong was asked to attend a workshop organized by the government, which was focused on getting across key messages from the state education authorities about the need to change the curriculum as a result of changes in society, and to ask school leaders to share their views and brainstorm on the strategies for implementing the new curriculum (OECD 2010):

> *The retreat usually started with a talk from a prominent community leader on how "society has changed". The curriculum development institute then outlined the curriculum reform, and each school delegation was asked to discuss their initial strategies for implementing it. The school groups then exchanged views.* (p. 104)

While the details of implementation were reportedly left to the schools, to preserve the principle of school autonomy and respect for the individual, it appears that in the case of Hong Kong, the state took charge of the transformation process, first by channeling discontent with the education system into a message calling for comprehensive change, and secondly, by building new structural frameworks for education and shepherding the schools into embracing and implementing its vision for reform.

3.3.4. FINLAND

Due to its consistently top performance on a major international student achievement study – PISA, implemented by OECD in over 60 countries and territories – Finland has long become the poster child of quality education. Its schools attracted scores of study tours by educators and policy makers from around the world, all seeking to uncover the secret of Finland's academic success, and the role that its state policy has played in achieving it. As with the other country cases above, I offer a brief overview of the state role in education in Finland.

Observers note that one of the key characteristics of the Finnish education system is its reliance on a comprehensive approach to schooling, and its elimination of the tracking, or ability grouping policies that were entrenched in its schools in the postwar period. While a difficult reform to push through, with resistance from the teachers of high-ability grammar schools, a comprehensive framework was adopted system-wide by the 1980's, and all ability grouping was abolished in primary and lower secondary grades (OECD 2010).

The development of the education sector during the early decades post World War II in Finland was marked by the remarkable growth of the Finnish capitalist and industrialist welfare state (Antikainen 1990; Houtsonen et al. 2010), with education increasingly seen as a necessary and required part of the provision of basic services by the state, and a way to equalize opportunities for social mobility across populations with unequal starting conditions. The elimination of tracking, coupled with an extension of compulsory schooling to include upper secondary education falls within this line of thinking. Antikainen (1990) argues that the development of the Finnish education system was "consciously managed" (p. 76), and education became an integral part of the political economy in the country.

Writing at the end of the 1980's, Antikainen (1990) recognizes the effort and the active role of the state in developing a national curriculum with a comprehensive learner centered philosophy, but points out to lingering inequalities in secondary grades, where the social status of the students often determines their educational experience. Unlike the unequivocally positive view voiced by OECD (2010), the author argues that policy makers chose to ignore the persistence of social inequities as an inconvenient reality. The author also argues that the system was influenced by neoliberal tendencies during this period, with a greater emphasis on individual school autonomy, and reduced role of the state with debates flaming about the legitimacy of central government to have authority over school-level decision making (ibid). Houtsonen et al. (2010) also notes the movement towards decentralization of school management and greater role of the schools, required to develop their own curriculum based on overall national guidelines in the early 1990's, which, according to the authors, was in large part accounted for by a the serious economic recession in the Finnish economy. School management was organized according to the principles of "goal orientation, evaluation, and accountability" (p. 601).

While opinions converge that the Finnish teachers and schools have a considerable degree of liberty in devising instructional approaches and management mechanisms at the school level, the central role of the state in building a teacher cadre with a shared understanding of the principles and the ideology of the education system. Teachers are held in high regard, are paid a good salary, and entry into the profession requires rigorous training in instructional methodology, child psychology, and advanced knowledge of subject matter content, for secondary school teachers. Admission into pre-service teacher training programs is competitive, and most of the graduating students enter the teaching profession. Furthermore, in addition to the technical mastery of teaching practice and subject matter knowledge, teachers are expected to ascribe to the principles of inclusion and internalize a feeling of responsibility for the achievement outcomes of all students in their classrooms (OECD 2010). In sum, the state builds a common frame of reference for educators throughout the system, through an intensive system of teacher preparation, and sets general parameters for teacher performance. Furthermore, a national accord on curriculum content helps ensure that all teachers throughout the country understand the requirements for their students performance, and are supported through the presence of instructional materials that address common content. This common frame of reference then allows it to set expectations and require school-level accountability without controlling the "how" of the teaching process at the classroom level. In short, the emphasis of the state is on heavy engagement "upfront", rather than intensive management of the education process.

3.3.5. CUBA THROUGH THE LENS OF CARNOY ET AL. (2007)

Finally, I introduce the case of Cuba, as it is described by Carnoy et al. (2007), in a comparative study of Latin American states conducted in the wake of the 1999 LLECE student achievement study administered by UNESCO. The study begins with a puzzle: an

economically disadvantaged Cuba substantially outperforms its wealthier neighbors on the continent, and demonstrates a remarkably equitable distribution of quality across its schools, regardless of whether they are located in central Havana or in the rural areas. Given the strong relationship between the wealth of a nation – measured by GDP per capita, as an example – and its successes on the educational front, one would not expect Cuba to top the list of Latin American states, showing a mean score more than one full standard deviation above the second in place.

Carnoy et al. (2007) begin with a broad quantitative analysis of the UNESCO/LLECE data across all of the participants of the assessment, and finds while socioeconomic context surrounding the Cuban education system – including the high overall level of formal education among parents, and near-zero unemployment – is important as a predictor of quality, it does not completely explain the difference between Cuba and the rest of the participants. Even with the same level of educational attainment, argues the author, Brazil and Chile, for example, would not have achieved the same results on LLECE. Therefore, according to Carnoy et al. (2007), an explanation must lie in the systemic features of Cuban education, and the ways that the state has shaped its role and relationship with the classroom. Using this as a working hypothesis, the researchers took an in-depth look at three Latin American education systems: Chile, Brazil, and Cuba, to explore whether the institutional characteristics of the state structures and policies, as well as the structure of their societies, has contributed to the surprising results on LLECE.

The authors found that the socioeconomic characteristics of the students are highly predictive of the types of schools they attend, and hence, their achievement outcomes, in the capitalist democracies, but not in Cuba. The school choice systems put in place in Chile in accordance with the theory of pluralism were intended to stimulate competition among

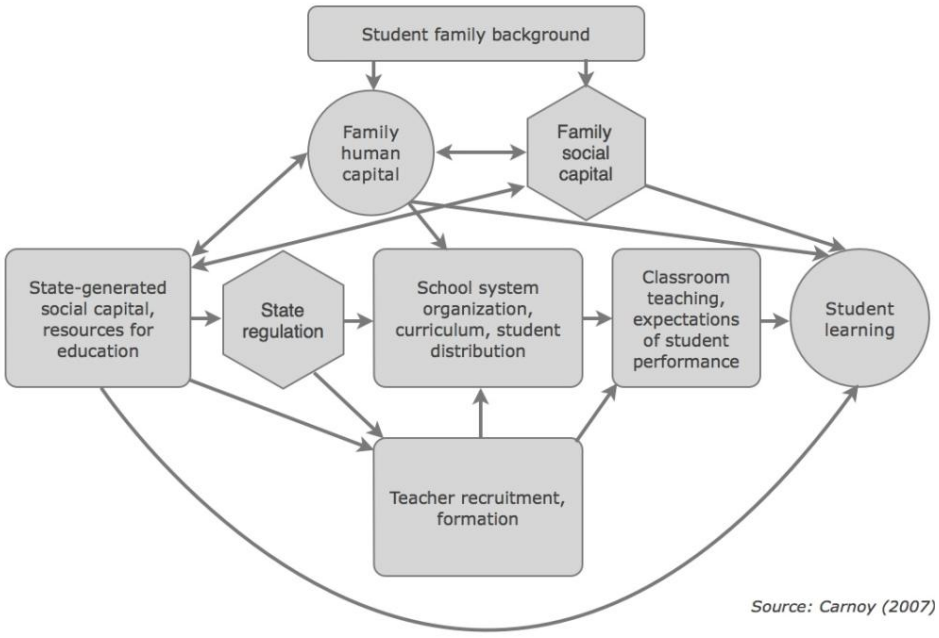
schools and consequently, increase the quality of schooling services, but in reality, they were unable to alter the inequality of educational opportunity already existing in the society. In both Chile and Brazil, the variance in average social class was very high *among* schools but not *within* schools, meaning that there was a high degree of concentration of students with peers of the same social class as themselves. In Cuba, there was also some differentiation between urban schools and rural schools in the level of education of their parents, but not to the extent as in the other two countries.

The emphasis on individualism at the expense of the State in Brazil and Chile was evident in the amount of autonomy given to the teachers and principals at the schools. In contrast to Cuba, in these two countries Carnoy et al. (2007) saw virtually no real mechanisms available to the State to ensure that the quality of education is consistent across the education system. In the socialist context of the Cuban state, according to the study, the delivery of services is organized hierarchically, with the lower levels of state education structures following orders from the central Ministry. As the education system of Cuba embraced child-centered pedagogy, the state pushed the delivery of child-centered methods through effectively established channels: teachers were trained in detail on how to teach student centered pedagogy, given lesson plans and guidelines, and offered regular assistance from school-level peer support networks. This may seem counterintuitive, as one may expect that student-centered methods would be given greater emphasis in more open societies and less controlling governments.

In Cuba's highly centralized education system, according to Carnoy et al. (2007) the vision for education quality was formulated and reflected in all of the relevant policy documents; policies were translated into practices which are delivered from the policy level to the school through direct mechanisms of implementation, with clear lines of

communication, responsibility, and accountability to the higher authorities, as well as to parents and communities. Since student-centered instruction was part of the state vision for achieving quality in education, such methodologies were delivered to teachers in a consistent and coherent way, through pre-service preparation, induction and mentoring mechanisms, ongoing in-service training, high-quality materials, and school-level instructional leadership. If quality was lagging in a given school, the state identified and addressed the quality issues, either with additional support or sanctions. In short, the state possessed strong central capacity to ensure close implementation of its goals in every classroom, equalizing educational opportunity and achieving high academic standards. Figure 3.1 presents the networks of relationships and linkages that impact student learning, as outlined in Carnoy (2007). This chart illustrates the multitude of direct and indirect ways in which the state affects achievement outcomes.

Figure 3.1. The network of relationships affecting student learning.



Based on these findings, the authors argue that this relationship between the State and education in the Cuban context can be characterized as "state-generated social capital: the state education structures in Cuba made it possible for less wealthy students, or those with less educated parents, to attain an education of similar quality to that of the more privileged students in the urban schools of Havana. These features were, according to Carnoy et al. (2007), the contributing factors in the success of Cuban students in the UNESCO assessment. By contrast, the decentralized states of Chile and Brazil lacked the strong coupling between state policy structures and the essential elements of the education system that impact quality learning: teacher capacity and school practice.

Another important defining characteristic of the Cuban educational context is that due to the structure of its economy, and the overwhelming presence of the state it all its sectors, teaching in Cuba – unlike in other countries in the region – is a prestigious and desirable profession with a relatively good pay. This makes it possible for the state to draw the "best and brightest" among high school graduates into teaching, and to effectively retain them through long-term careers. In some ways, one may argue that the state in Cuba has "an easier time" maintaining the quality of instruction, due to this ability to attract talent into teaching by sheer absence of other career options, and that therefore, the starting conditions before the national education policy is engaged are very different.

Finally, Carnoy et al. (2007) listed "tight coordination with existing curriculum" as one of the greatest strengths of the Cuban education system, and its advantage over other Latin American countries. They argue that the autonomy of teacher education programs in Brazilian and Chilean universities comes at the cost of the distance it creates from practice, and the lack of opportunity for teachers to master the content they are expected to teach

upon graduating. By contrast, in Cuba, the main goal of teacher pre-service institutions is to train teachers how to reach national curricular objectives.

3.4. CONCLUSION

In this chapter, I reviewed the contributions of several authors along two main threads of literature: 1) the theory and practice of student-centered instruction; and 2) the examination of the role of the state in effective education systems. In the first part of this chapter, I reviewed the major theoretical contributions to the development of student-centered teaching in the West and in the former Soviet Union, and looked at a few studies that attempt to quantify the value of student-centered instruction for student achievement outcomes. In the second part, I briefly discussed the cases of a few top performers on international tests –Singapore, Japan, Hong Kong, Finland, and Cuba. This review, although by no means exhaustive, offers a few insights that led me to the development of a theoretical framework for this dissertation (Chapter 4), and influenced the design of the methodology for my research (Chapter 5).

One, it is clear that student-centered instruction is not a set-in-stone package of instructional practices, but an umbrella term for methods that are designed to generate student participation in his/her own learning. The theory of learning behind this general concept was enriched by contributions of authors from the point of view of child socialization, child psychology, mind and brain development, as well as emotional well-being of the child. Various aspects of teaching and learning were emphasized, including group work, project-based learning, the use of experiments and demonstrations, reflection and discussion, observation and relation to the student's everyday life – but no method was singled out as the most beneficial for learning.

Results of student-centered teaching on achievement outcomes are far from conclusive. Smaller-scale studies of targeted interventions claim to show large effects on learning, but the selection of students into such studies cannot be completely random, making it difficult to ascertain the generalizability of their findings. Large studies using the data from international achievement studies have generally struggled to find any evidence of relationship between classroom methods and achievement outcomes.

As we looked at the level of the state, in the case studies I discussed above, a curious similarity emerges: student-centered instruction and an "emphasis on learning" remarkably is the theme of all state-sponsored education quality improvement efforts in these countries. The pathways to the adoption of this theme are different, as are the implementation mechanisms: in some cases, the state generates the frames of reference and follows through on each step of implementation, including perception management to ensure successful internalization of its vision; in other cases, the state responds to pressure from the society, builds a framework for implementation, and ensures the teachers have the capacity and the common understanding to implement it, but stays away from control of classroom-level activities.

However, in all of the cases I reviewed, there was a common frame of reference, a shared understanding of what education was about, what its goals were, what role the government played in defining the parameters of public education, and what was expected of each school in order to accomplish them. This common frame of reference was in every case built and supported by the state. To the extent that all parties converged on this shared understanding, there was coherence in the implementation of new curricula, the creation of more equitable school structures, and the types of outcomes that were expected of the students. In almost each case, the society and the state went through a series of pendulum

swings – from rigor and content-centered emphasis, to process-oriented practices where the learning process per se was the goal, to back-to-basics movements critical of the easing of school-related pressure. However, once a common frame of reference was shaped, the institutionalization of effective practices could have been top-down or grassroots –up.

The Cuban case presents an example relevant for the examination of the education system of centralized post-Soviet state – and particularly, Kazakhstan, both due to their structural- historical commonalities (Soviet models of education were often replicated in countries under its sphere of influence, of which Cuba was a prime example), as well as due to the level of concentration of power in the hands of the central government and the mechanisms available to the state in curbing the choices of individuals for a perceived greater good. This case also offers a strong theoretical argument in favor of a strong role of the state in shaping education – albeit, as I will show, with caveats regarding the specific leadership roles a centralized state can play in strengthening instructional practices. In Chapter 4, I weave the Carnoy et al. (2007) argument into the theoretical framework for a conceptual understanding of the context for instructional transformation and the use of student-centered instruction in Kazakhstan, along with the research hypothesis setting the expectation for the relationship between student-centered instruction and achievement. Chapter 5 presents the mixed-methods approach designed to test these hypotheses.

CHAPTER 4.
THEORETICAL FRAMEWORK

4.1. INTRODUCTION

While student-centered pedagogy marks a considerable change from the traditional direct instruction that is the default mode in most post-Soviet school systems, it is by no means a new phenomenon in pedagogy. One could argue that its origins go back as far as ancient Greece, when Socrates engaged his students in conversation and reflection about abstract concepts, and encouraged them to relate knowledge to their daily lives (Common, 1994). In John Dewey's schools, children learned through work, and were responsible for completing their work projects as part of learning subject matter content. In the Soviet Union, the seminal works of Lev Vygotsky influenced child psychologists and educators around the world through his thesis of the zones of proximal development, which in essence, encourage the teacher to serve as a facilitator and supporter of the child as she seeks out and acquires new knowledge and skills. Yet another famous Russian educator, Sukhomlinsky, argued that an individualized approach to the student, the encouragement to learn through play and the creation of a warm, friendly atmosphere where discussion was an integral part of the lesson were crucial and yet largely missing elements in Soviet schools. All of these scholars have influenced the concept of student-centered instruction as it is understood today. These and other, more contemporary authors are reviewed in greater detail in Chapter 3.

However, despite the long history of student-centered instruction as a concept and approach to teaching, as I showed in Chapter 3, its relationship with student achievement is not conclusive or definitive. Studies show varying results, and it appears that while in some settings, there may be clear positive gains in learning associated with a particular student-

centered methodology, such positive effects are extremely difficult to discern when one examines a broader sample of students, and accounts for the variation in the background characteristics of students, schools, and teachers correlated with their choices of methodology (for an overview of studies see Chapter 3).

In this chapter, I outline my hypothesis of the impact of student-centered instruction, by drawing the connections between the teaching approach and the context for its implementation. In essence, I argue that student-centered methods as an instructional approach can be effective in improving student achievement outcomes if they are institutionalized and implemented in a consistent manner by qualified and competent teachers across a critical mass of schools. I hypothesize further that there can be alternative pathways for such institutionalization, and the choice of the pathway depends on the context of the education system, and consequently, on the structure of the relationship between the state and the individual in education. Based on this distinction, I propose a framework for understanding the place of student-centered instruction in the post-Soviet state, and set out the conditions under which it can be effective in that context, by drawing parallels between the post-Soviet state and the state of modern Cuba, as described by Carnoy et al (2007). This hypothesis will then be tested first through a quantitative analysis of the linkages between student-centered instruction and achievement (Chapter 6) and the further explored though qualitative data on the realities of educational development in Kazakhstan (Chapter 7).

4.2. EFFECTIVENESS OF STUDENT-CENTERED INSTRUCTION

While there are many elements that compose an effective, well-functioning education system, in this dissertation I focus on classroom instruction – and specifically,

instruction using student-centered methods, as distinct from the traditional chalk-and-talk, direct instruction method that is largely the default mode in most schools throughout the post-Soviet space and beyond. Using the literature on teaching and learning that I discussed in Chapter 3, I form several assumptions about the relationship between student-centered methods and student achievement, which I will then test using quantitative methods and further explore through analysis of qualitative contextual information.

4.2.1. RESEARCH HYPOTHESIS AND EXPECTED OUTCOMES

Because there is no set definition of what combination of methods would constitute a student-centered learning environment, the core characteristic that threads through the literature on teaching and learning is active participation of the student in classroom activities, and his active physical and intellectual engagement in digesting new information and applying to the task at hand. This would involve engaging in group discussions; group work on projects; individual assignments requiring a critical and creative approach to a problem; reflection on various aspects of a given problem; unhindered, open dialogue with the teacher on the subject matter material the student is learning; conscious observation of natural phenomena and application of acquired knowledge to their understanding; and the application of the scientific method – through generating hypotheses and their testing in an experimental setting. Further, since memorization and unassisted completion of problem sets given out by the teachers in class or for homework are undoubtedly essential and inseparable components of classroom teaching, I hypothesize that these methods do not disappear in student-centered environments, but serve as a necessary foundation for the successful implementation of the more active and participatory learning methods.

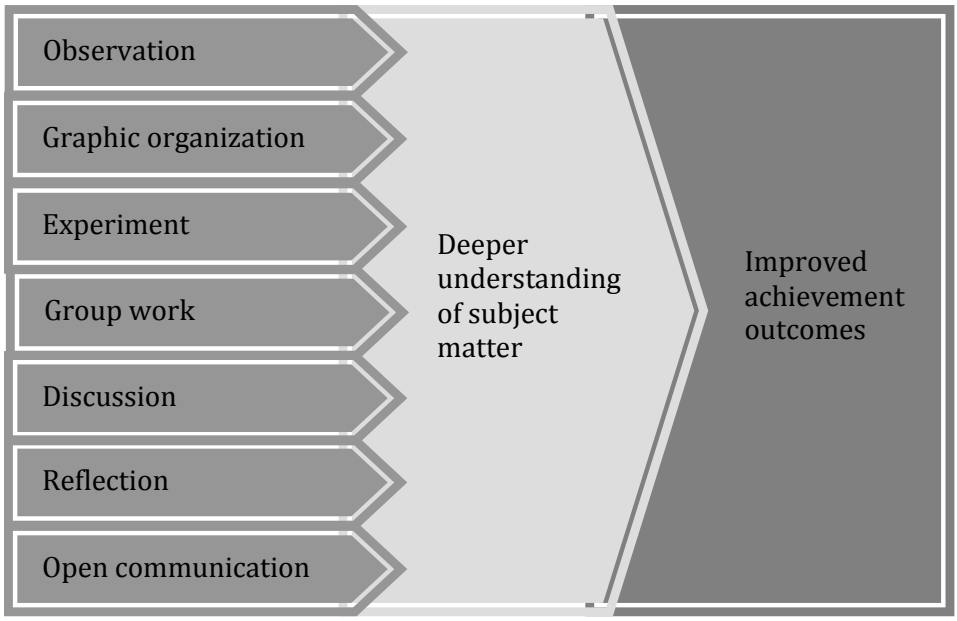
The learning theory behind these methods is that through active participation in the learning process, the student is able to "make sense" of the subject matter, and gain a deeper understanding of its impact on the world around her. By engaging all of their senses – eyesight (observation, reading), hearing (teacher-led discussion), motion (taking measurements, conducting experiments, learning though play), and social interaction (group work, group discussion) – students form a comprehensive picture of how each element of what they are learning fits into the framework of their previous knowledge, and where more knowledge is necessary to form a complete understanding of a given phenomenon. This deeper knowledge leads to the development of an intuitive ability to address problems, tasks, and assignments, not only in the specific subject where these methods are applied, but more generally.

In this dissertation, however, I am interested in a more narrow aspect of learning: academic achievement, or the ability to master the curriculum content in a given cognitive domain. I want to know whether there is evidence that student-centered methods in teaching – which are, as I noted above, prominently stated as the new approach to instruction in Kazakhstan – could have influenced student achievement in mathematics and science, measured by an objective, standardized instrument. By means of comparison, I seek evidence that students who received more exposure to student-centered instructional methods in their mathematics and science lessons did substantively and significantly better than their peers in purely traditional settings.

Given the complexity of capturing the extent and intensity of a student-centered environment, I confine my focus to the measurable aspects of this instructional approach, with the understanding that such measurable aspects are but proxies of the true nature and quality of the student-teacher interaction that form the essence of the learning process.

This involves examining the specific components within the general student-centered approach, and namely, the presence of such activities as *observation*, *scientific experimentation*, *group work*, *group discussion*, and *reflection* and *open communication* with the teacher about the subject of study. Figure 4.1 shows the conceptual framework of student-centered instruction and its expected impact on achievement. I examine whether the hypothesis that extended experience with such activities in the classroom positively affects achievement may conceivably be true, all other relevant factors holding constant. Extended experience is defined as continuous, regular exposure to these teaching activities in math and science lessons, as opposed to haphazard exposure (as would be the case if the teacher starts with a method and abandons it, or uses it only on specific occasions).

Figure 4.1. Conceptual framework of the impact of student-centered instruction on achievement



Source: Author, based on literature (see Chapter 2)

There are several potential ways of looking at the relationship between student-centered methods and student achievement. One hypothesis is that each practice is instrumental as an individual independent element, capable of affecting achievement

outcomes on its own, over and above the associations generated by other practices. Another hypothesis is that any student-centered practice makes no difference if it is not implemented as part of a comprehensive approach, and in order to expect any learning results, one must form some gauge of a collection of practices as a single unified teaching approach and establish its association with achievement, as a whole. Yet another hypothesis can be that not all practices are equally important, and some combination of methods and their intensity is sufficient to form the basis of a student-centered environment. I test all of these hypotheses in this dissertation, and employ a variety of approaches to measuring student-centered instructional practices: as individual survey items, for math and science separately, as binary variables that act as composite indicators of whether or not student-centered instruction was offered to students in a given classroom.

4.2.2. ALTERNATIVE HYPOTHESES

As I explained in detail above, my working hypothesis for this study is the positive relationship between student-centered instruction and student achievement, holding constant all other factors. The two alternative outcomes for my analysis, then, are: 1) no relationship between student-centered instruction and achievement (null hypothesis); and 2) a negative relationship, which would mean that students receiving greater exposure to student-centered methods perform substantially worse than their peers in traditional learning environments. There are several potential reasons why either of these alternative hypotheses might be true.

First, it may be the case that student-centered methods have no impact on cognitive achievement, and instead affect the students' social and emotional development, rather than

their mastery of any concrete subject matter. Such an approach may, for example, enhance the students' ability to address problems related to their daily lives and cope with challenges not related to their academic performance. Because the comparison is between learning environments with student-centered instructional methods and purely traditional direct instruction, achievement levels may be attributable entirely to chalk-and-talk methods, with no additional value added by the use of student-centered activities. The null hypothesis would suggest that if one decides to use student-centered instruction, *other* outcomes, perhaps non-cognitive outcomes, should be expected as a result. This would not mean, however, that if student-centered instruction were compared with no instruction at all, that there would be no impact.

Secondly, student-centered instruction may be considerably more difficult for teachers to carry out effectively than traditional instruction. Because of the emphasis on the primary role of the teacher as facilitator, teachers may find it difficult to strike a balance between taking responsibility for the learning process in their classroom, and letting the students actively participate and discover knowledge in an open learning environments. A misinterpretation of student-centered approach as that which allows the teacher to withdraw completely from delivering curricular content and shift that burden entirely on the students, for example, may lead to not just a lack of effect, but to a negative result where students learn less than they would otherwise have in a regular classroom. In such circumstances, my second alternative hypothesis – that of negative impact – may come true.

Finally, there may be circumstances relevant to a specific educational system context – in this case, that of a post-Soviet state – that would make the null hypothesis more plausible than a hypothesis of a positive relationship. There may be factors in the general fabric of society that determine the effectiveness or lack thereof, resulting from student-

centered, open learning methods in math and science. A change in methodology may require time to take root and begin to bring about expected results. In such environments, it may also require substantial support from the state, and a continuous development of teaching capacity, through training and professional development of teachers and school administrators. This hypothesis can and will be explored in this dissertation through a qualitative data collection and analysis examining the context of Kazakhstan as a post-Soviet highly centralized state. The following sections present the argument that shapes my approach to analyzing the role of systemic factors in the effectiveness of student-centered teaching.

4.3. THE CUBAN SUCCESS

Carnoy et al (2007) presents a good case of a system-wide implementation of effective instruction in primary and secondary school; and while it is primarily a case of a successfully functioning tightly-coupled educational system, incidentally student-centered methods occupy a prominent place in the "moral imperative" for quality education in the Cuban state. The common approach to instruction starts from pre-service teacher training, where student-teachers are first initiated into the profession, and is reinforced through continuous in-service professional development and on-site instructional leadership and support at the schools. Child-centered pedagogy is *the* common approach to the delivery of the national curriculum, which serves as the foundation for pre-service and in-service training, as well as the core content of instructional support materials. Through professional links between pre-service training colleges, schools, and the governing structures of the Ministry of Education, educators share a common understanding of the vision and goals of the country's educational development, and a common framework of what constitutes effective instruction – in terms of its actual operationalization, as well as

its expected outcomes, potential challenges, and available support resources. The authors make a strong argument that it is this tightness of connections between the different elements of the education system, as well as the shared understanding of curricular and instructional goals, that played a crucial role in Cuba's academic success despite its relatively modest economic achievements.

However, the descriptive features of Cuba's education system are, undoubtedly, strongly linked to the country's political and socioeconomic conditions. As one of the few remaining Communist states, Cuba enjoys a low level of social inequality, particularly in comparison with its neighbors in Latin America; a low level of unemployment, and a near-universal attainment of at least secondary education among its adult population. Furthermore, due to the closed nature of its economy, Cuba was able to preserve a relatively high status of jobs in the education sector, which has meant that graduates of teacher training colleges immediately joined the ranks of teachers in the nation's schools, replacing retiring teachers and preserving the continuity of instruction.

Furthermore, the political characteristics of Cuba as a one-party state with a strong, authoritarian leader at the top, have influenced the nature of the relationship between the state and the individual in education. Carnoy et al. (2007) point out to the restrictions placed on individual choice and preference in education: teachers are not free to choose the content or method of instruction, but must follow the official national guidelines, and work in collaboration with their peers to ensure that a uniformly high level of quality of their teaching. Parents are encouraged to support schools, but the education of the nation's children and their attainment of high educational standards is the responsibility of the school, and schools are held accountable for the performance of all students.

However, while Carnoy et al. (2007) acknowledge that these underlying conditions have played a role in the success of the education system in Cuba, they argue that they are not absolutely necessary for the state to assume the same level of responsibility for the country's educational outcomes. It is the existence of the common "moral imperative" for education, the ability of all actors to work collaboratively, the commitment to take responsibility for overcoming the influence of social inequality, and the sharing of a common frame of reference in instructional and curricular goals among educators, that are, according to the authors, the crucial ingredients of broad-based academic success.

4.4. PATHWAYS TO INSTITUTIONALIZATION

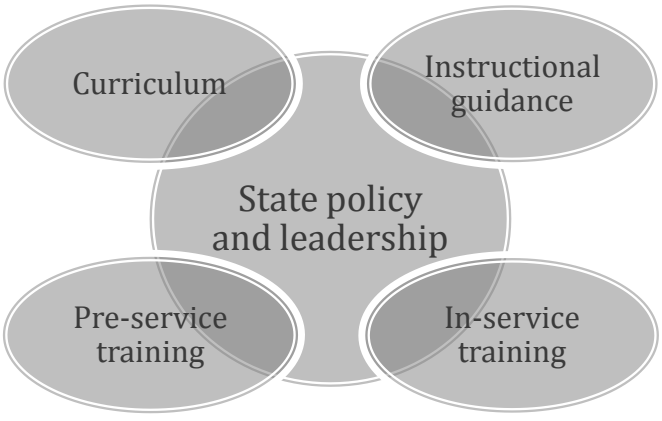
Leaving the political and socioeconomic conditions aside, I hypothesize that effective student-centered instruction accounts for a large portion of Cuba's academic advantage over its neighbors in Latin America. I hypothesize further, based on Carnoy et al. (2007) that instruction in general has been highly effective in Cuba due in large part to the institutionalization of innovative practices and tight linkages between the different elements of the education system, as a result of strong involvement and oversight on the part of the state. The state ensured that an adequate level of resources and human capital were available to support and oversee teachers, from the point of their induction into the profession until their retirement, and that effective pedagogy was evaluated, adopted, and disseminated. Teachers were never positioned as individual professionals, but as agents within a large system, with its guidelines, support networks, and a common vision.

I define institutionalization as the formation of a common referential system that shapes the understanding of what excellence in pedagogy really means, and what demands it places on the efforts of a given teacher in a classroom. Dissemination and subsequent

institutionalization can be achieved through different avenues, equally legitimate and capable of bringing about success, but requiring different sets of conditions.

One avenue is school and teacher networks and professional organizations, where teachers can share their innovative practices and collectively form perspectives on "what works" in education – including methods of instruction, classroom management, and support of students from disadvantaged families. Peer support networks have been researched and generally seen as beneficial for teacher professional development and hence, for the improvement of teacher capacity for instructional excellence. The success of this avenue to the dissemination and institutionalization of excellence in pedagogy hinges on: (a) strong foundation of knowledge and skills in classroom management and effective pedagogy among a critical mass of teachers; and (b) the presence of an active civil society environment that allows teachers to form professional networks, share ideas and best practices, and influence the formation of the national curriculum, instructional choices for pre-service and in-service training, and national education policy for the institutionalization of widely adopted approaches.

Figure 4.2. Role of the state in a centralized education system



Source: Author

Another pathway to the institutionalization of innovative and effective methods of instruction is through active engagement of the state. In this scenario, the state both directs the formation of policy that provides the conditions and incentives for excellence in instruction and school management, and follows through each step of its implementation by all teachers across the system – who in this scenario are employed by the state and therefore are under its direct control. The state actively engages in research and evaluation to shape a common understanding of which instructional practices are beneficial for educational quality, and has a strong central capacity that transforms that knowledge into education policy. The policy is then aligned with national curriculum, and disseminated through pre-service and in-service teacher and administrator training, such that it forms a common, shared narrative of effective instruction. The state ensures that all actors in the education system, including students and teachers, subscribe to this narrative and do their part in implementing it into reality. The success of this scenario lies in: a) strong capacity of the state to build policy and connect pre-service teacher training, curriculum development, and in-service professional development into one mechanism; and b) the willingness of agents within the system, as well as non-state actors, such as students and parents, to follow along, either as a result of a restricted set of choices, as in Cuba, or their genuine trust in the ability of the state to deliver quality.

4.5. QUALITY OF INSTRUCTION IN THE POST-SOVIET STATE: THE SCENARIO FOR KAZAKHSTAN

Thinking of education reform – and instructional reform in particular – in the post-Soviet state, we can expect that the second scenario would be more likely to unfold in the more centralized environments, such as the countries of Central Asia, and certainly Kazakhstan. Indeed, the formation of grassroots professional networks of educators with

substantial say in shaping the common frame of reference for education is less plausible in authoritarian environments where civil society groups are generally seen as a potential threat to the ruling regime. While innovation can still occur in haphazard and heterogeneous ways at the grassroots level in such systems, such heterogeneity is more likely a result of a lack of insufficiently strong capacity on the part of the state, rather than its willingness to absorb newly developing forms of teaching. The state in such a political system would seek to extend its control over the content and method of instruction, through all available means. Furthermore, it is more likely that due to the history of centralized authority, schools in such environments will look to the state to endorse acceptable forms of curricular content, instructional approaches, and other aspects of school management.

As I noted above, the crucial assumption necessary for the success of the state-driven scenario is the capacity of the central state education bodies to generate and sustain effective models of instruction, starting from initial teacher training, to instructional materials, to professional development for teachers and principals, to instructional support at the schools, and finally to assessment. Importantly, the state must also maintain a sufficient level of resources, so that it could ensure that school infrastructure is in reasonable condition, prevent wastage of pre-service training, and keep teacher attrition to a minimum. The advantage of a centralized state – which Cuba used to the fullest – is its ability to restrict individual choices and enforce compliance with the roles and responsibilities prescribed by the state. States like Cuba achieve such compliance by creating a political, economic, and social environment where citizens feel the omnipresence of the state, and are cultured to expect substantial limitations to their individual freedoms. In this dissertation, I explore to what extent the same dynamic, and more importantly, the

same level of technical and managerial capacity is present in the case of the post-Soviet centralized and authoritarian state: Kazakhstan.

If my hypothesis is true, one can expect Kazakhstan to use its central authority in education to its maximum, in order to rebuild and reconfigure its education system as that of an independent state, rather than a former Soviet republic. There is evidence in the official rhetoric that the ambition is present to make a definitive shift towards "new ways" of teaching and learning, and "shifting the focus of the educational process from the teacher to the student" (MOE 2005). In this dissertation, I narrow my focus to student-centered instruction as the core element of this shift, exploring both whether it shows any evidence of association with achievement, on one hand, and whether the state is the core actor in instructional reform, equipping teachers with knowledge and skills to change the dynamics in the classroom, and ensuring the cohesion of elements in the education system.

If, however, the key assumption of central capacity is not met, then the centralized state is unable to generate a coherent approach to quality in education, and hence will be more likely to direct its efforts at maintaining the status quo, or engage in *double-talk*, adopting reform and innovation on a discursive level without taking it to the classroom (Steiner-Khamsi & Stolpe, 2006). An alternative path would be to let innovation and best practices in effective teaching grow from within the system, or to allow third parties, such as nongovernmental organizations with a mandate in education, to bring in expertise and innovative ideas for teaching, and gradually assimilate and institutionalize the most effective ones system-wide. Professional communities of educators would provide opportunities for dissemination, and offer opportunities for teachers to learn from their peers. However, as I noted above, authoritarian political systems may find it difficult to let their education systems form professional communities, or let outside parties build

constituencies among educators for their models of effective teaching. This inability or lack of willingness at central levels to identify and assimilate successful practices emerging from the engagement of various non-state actors at the grassroots is a weakness of centralized education.

Whatever the actions of the state are in the absence of central capacity for instructional leadership, they are directed at strengthening state control of the education system, be it through visible injection of material resources, or structural reorganization of various elements within the system, or through the development of external accountability mechanisms intended to demonstrate the state's ability to take responsibility for student achievement outcomes.

Using the hypotheses outlined in this chapter, I embark on a mixed methods analysis of classroom-level quantitative data and contextual qualitative information from interviews. The next chapter presents the rationale for the choice of methodology, as well as a detailed description of both the quantitative and qualitative methods.

CHAPTER 5.
METHODOLOGY

5.1. INTRODUCTION

Because this dissertation pursues two levels of analysis – the level of individual student achievement outcomes and the role of the state in education – two complementary approaches are used to address the research questions outlined in Chapter 1. In order to examine the associations between instructional practices and various confounding background variables and student achievement, I apply quantitative methods, with particular emphasis on exploratory strategies such as factor analysis and ordinary least squares regression, and quasi-experimental methods such as propensity score matching. For an in-depth examination of the role of the state in setting instructional practices and attaining higher quality achievement outcomes, I turn to qualitative case study research, relying on key informant interviews, focus group discussions, and document review. In this chapter, I provide a brief rationale for this choice of methodology, followed by a detailed description of the methods. Because the bulk of my analysis lies in the quantitative elements of the dissertation, the description of quantitative methods occupies the major part of this chapter.

5.2. RATIONALE FOR MIXED METHODS ANALYSIS

The choice of methods in this dissertation was determined by the research questions that I sought to address, and the nature of the subject of study in addressing these questions. First, in order to argue more or less convincingly about the benefits (or lack thereof) of a given teaching approach across an education system, one must base such arguments on representative sample of beneficiaries of this approach (in this case, primary

school students), as well as a sufficient number of counterfactual cases. The unit of analysis at this level is the individual student and his/her achievement outcomes. It goes without question that individual student characteristics, their learning styles, motivation and aptitude compose a great deal of what determines their academic achievement. Education research, in essence, looks for marginal differences that can affect student performance – particularly when it comes to manipulable elements such as educational policy, teacher training, or instructional methods. It is therefore imperative, in research examining broad system-wide effects of such elements to minimize the influence of background factors and selection bias, which is best attained by quantitative analysis of survey data. The TIMSS 2007 sample provides just such an opportunity, with its rigorous two-tier sampling structure, rigid requirements for the definitions of the student population to be tested, and high response rates across all of the instruments – from the test itself to the accompanying background questionnaires for students, teachers, and school administrators. I therefore leverage the richness of this dataset extensively, using data from all of the questionnaires, and looking for relationships between the nearly 4,000 student responses on the survey about their classroom experiences within a country, and their achievement outcomes in mathematics and science. This dataset allows me to isolate the contribution of extraneous variables on achievement, and on the likelihood of a given instructional practice being offered to a class of students. In sum, quantitative analysis makes it possible to draw conclusions about the presence or lack of positive relationship between student-centered instruction and achievement across the education system in Kazakhstan, given the observed characteristics of students and their learning environments.

The qualitative analysis, on the other hand, compensates for the lack of depth in quantitative data, and provides context for the interpretation of findings from statistical

hypothesis testing. For example, the TIMSS background questionnaires offer no information on the quality of the application of instructional methods by teachers, their intensity and appropriateness for the tasks in the classroom. They also do not convey the teachers' level of comfort with various methodologies, their level of understanding as well as beliefs about the overall effectiveness of various methods, and their rationales for the choices of methods in their work with students. One also cannot perceive, just based on the TIMSS data, the broader picture of the instructional environment in Kazakhstan, and the pressures and incentives that educators and students are facing that inevitably influence achievement outcomes. The qualitative case study seeks to fill this gap in data through interviews with teachers and school administrators.

The qualitative case study also focuses on the state as the unit of analysis, looking at how centralization and tight coupling benefit or hinder its ability to bring about, and more importantly, to sustain educational quality. The contextual information gathered from documents and interviews offers a deeper look at the post-Soviet state, its weaknesses and strengths, as well as the narrative it builds around its national education system. Using the Carnoy (2007) argument on the effectiveness of the centralized state with a strong imperative for equal educational opportunity as a lens for looking at the highly centralized education system in Kazakhstan, I use the qualitative data to explore the nature of the relationship between the state and the primary school classroom. In the interviews and documents, I look for indications of the level of coupling between the central and local elements of the system, I seek evidence of a conscious choice of instructional methodology, and of leadership on the part of the state in holding the system together and maintaining a high level of quality. I also examine *how* the government of Kazakhstan translates its vision,

or narrative, into action in education, and what pathways it uses in stimulating quality of instruction in mathematics and science.

In sum, a mixed methods approach is the optimal strategy for addressing the goals of my dissertation project. Rigorous quantitative analysis of TIMSS data provides an unbiased look at the association between student-centered instructional methodology and various aspects of achievement, and allows me to generalize my findings to the entire population of Kazakhstan's fourth graders. It also offers a glimpse of a comparison of coefficients of association found in Kazakhstan with other countries, using exactly the same metric. The qualitative analysis helps to provide context for understanding these coefficients, and to test the hypothesized connection between the centralized state and the delivery of classroom instruction in core subjects.

5.3. QUANTITATIVE ANALYSIS OF TIMSS DATA

5.3.1. GENERAL DATA DESCRIPTION

As described above, in order to examine the association between student-centered methodology and student achievement in math and science, I used the TIMSS 2007 fourth grade assessment data. TIMSS, or Trends in Mathematics and Science Study, is administered by the International TIMSS and PIRLS Study Center at Boston College, under the auspices of the International Association for the Evaluation of Educational Achievement (IEA), every four years since 1995. Participation of post-Soviet countries in TIMSS, as well as in other international assessments, has been relatively low, with only Russia and Latvia regularly appearing on international league tables. However, increased interest in measurement and evaluation in education in general, and in measuring quality of learning in a comparative perspective, has prompted an increase in participation in the 2007 round of TIMSS, with as

many as seven former Soviet republics taking place in the assessment – including, for the first time, the Republic of Kazakhstan.

In Kazakhstan, like in most countries participating in TIMSS, the sampling was done in a two-tiered structure, where schools were first randomly sampled from a nationally defined population, and then one to two intact classrooms were sampled within each of the schools. The nationally defined population in Kazakhstan included all Kazakh and Russian language schools, and the test was administered in these two languages. The sample consists of 3,990 students in 141 schools, including 2,407 students tested in Russian and 1,583 tested in the Kazakh language. While the geographic location of their schools is not provided in the dataset, the sample is drawn from across various types of communities, starting from the very small villages of a few thousand residents, to large cities of 500 thousand and more residents.

While most of the in-depth quantitative analysis focused on the Kazakhstan's data, I expand the breadth of the study and put Kazakhstan in a comparative perspective, by exploring the TIMSS data from several other countries in post-Soviet Eurasia, namely Armenia, Georgia, Latvia, Lithuania, Russian Federation, and Ukraine. The models I fit on these datasets provide a snapshot of their learning environments and the relative contribution of various factors, including teaching methodologies, to the variation in student performance in mathematics and science. Table 5.1 provides the basic description of the TIMSS student samples in these countries, with the characteristics of the Kazakhstan's dataset highlighted in light grey.

The two-tier sampling structure means that the individual units of analysis – the students – are not, in the purest sense, independent observations, but are clustered in

groups. Students enrolled in the same classes are more likely to have similar backgrounds, and share unobservable characteristics such as motivation, natural aptitude in certain subjects, and work ethic. Because this rends implausible one of the key assumptions of classical regression – independence of errors – my analyses account for this feature of the sample, either through the clustering of standard errors, or through explicit multilevel modeling.

Table 5.1. Basic characteristics of the TIMSS country samples.

	ARM	GEO	KAZ	LAT	LIT	RUS	UKR
Sample size	4079	4108	3990	3908	3980	4464	4292
Proportion of female students	48%	48%	51%	50%	49%	50%	49%
Average age	10.6	10.1	10.6	11.0	10.8	10.8	10.3
Born in country	71%	85%	92%	93%	92%	94%	87%
Have at least 2 bookcases of books at home	32%	34%	18%	32%	16%	29%	23%
Have a computer at home	41%	39%	36%	81%	80%	59%	45%
Have a dictionary at home	75%	71%	77%	86%	81%	87%	90%
Have internet connection at home	22%	20%	21%	65%	62%	32%	26%
Live in large cities (>100,000 residents)	33%	43%	44%	27%	39%	51%	45%
Number of schools in sample	148	144	141	146	156	206	144
Average school size (enrollment)	538.9	568.4	772.8	509.3	607.4	578.7	583.4
Number of teachers in sample	228	405	175	339	283	268	192
Female teachers	88%	98%	93%	100%	98%	99%	100%
Certified teachers	92%	98%	100%		100%	99%	100%
Teachers with higher education (ISCED 5-6)	99%	99%	69%	97%	95%	73%	85%
Note: percentages based on unweighted samples.							

Along with assessments of mathematics and science, TIMSS administered a set of background surveys for students, teachers, and school administrators, which provide a wealth of information about the learning environments in the schools. I draw on these surveys extensively to construct both the treatment variables and covariates that are

known to affect student achievement, in an effort to first examine what selection bias is present in the data, and then to isolate extraneous influences of factors outside instruction, to examine relationship purely between instructional methods and student achievement. The next section describes the method of defining the treatment variable: student-centered instruction, as well as the variables reflective of traditional instruction, examined as a counter-hypothesis in this dissertation.

5.3.2. DEFINITION OF TREATMENT

As the problem statement of this dissertation emphasizes, there is no set or pre-determined package of student-centered instructional practices to serve as the subject of this research. What methods of instruction "count" as student-centered is a loose definition, based on the literature on democratic education, active learning, and critical thinking teaching methods, and like almost everything in education, student-centered teaching is subject to the interpretations and varying perceptions of both those who practice them – the teachers, and those who consume them – students and their parents. With large-scale surveys such as TIMSS, drawing nationwide samples, the ability of a researcher to identify a particular package of instruction and measure its impact is even more prone to error, because survey items by definition cannot refer to standardized policies and practices of a given education system, but must be general enough to be relevant across borders.

I attempt to generally define who can be roughly considered as having been taught using student-centered methods, and I use this definition to examine whether that group of students had significantly different achievement outcomes, holding constant all relevant background factors. It must be emphasized that this definition is arbitrary, and may be altered to be more or less stringent, although such alterations would not improve the level

of noise in the variable, as long as it is based on the same survey items. The following explains the coding of the "treatment", or student-centered instruction in mathematics.

5.3.2.1. ORIGINAL ITEMS

The items for the construction of the composite measure of student-centered instruction were taken from the TIMSS student survey questionnaire. The survey items asked the students to reflect on how often they were asked to perform certain kinds of learning activities in their mathematics and science lessons. These included the most common teaching methods, such as memorization and working through math and science problems in their notebooks, to the more active methods, such as group work and planning and conducting experiments in science. Table 5.2 displays all of the items from the student questionnaires that provide information on the teaching methods used in the math and science classes. As the table demonstrates, these are but a few features of the learning environment, and by no means constitute an exhaustive list of teaching methods. Furthermore, only the frequency, and not the appropriateness, depth, or the extent to which the students understood or enjoyed the learning activity were measured. Therefore, a lot remains outside of the realm of this study, which may account for the ambiguity of the results. At the same time, the breadth and consistency of information provided by the TIMSS student survey surpasses all available data on the teaching and learning practices in Kazakhstan – and in fact, it may be the only reliable source of information on classroom practices in the country to date.

Prior to the construction of the treatment variables (the two specifications of the treatment are described at length below), I corrected the data for noise by pooling student responses on items listed in Table 5.2 at the teacher level, and obtaining a mean value for

the frequency of a given activity for a group of students. In most cases, one teacher taught an entire class of students, and therefore, the pooling of student responses at the teacher level also coincided with class grouping. This is a valid approach to capturing instructional practice, since it is unlikely that different instructional strategies would be applied to different students by the same teacher within the same class. Pooling responses in this manner minimizes the noise generated by the random variation in student perceptions of the teaching practice, and offers a more consistent and stable proxy of the actual instructional methodology used in a given class.

Table 5.2. Teaching practices measured by TIMSS in 4th grade mathematics and science.

Math: How often do you do these things in your mathematics lessons?	***Science: In school, how often do you do these things?***
- **I practice adding, subtracting, multiplying, and dividing without using a calculator.**	- I look at something like the weather or plant growing and write down what I see
- **I work on fractions and decimals**	- I watch the teacher do a science experiment
- **I measure things in the classroom and around the school**	- I design or plan a science experiment or investigation
- **I make tables, charts, or graphs**	- I do a science experiment or investigation
- **I learn about shapes such as circles, triangles, rectangles, and cubes**	- I work with other students in a small group on a science experiment or investigation
- **I memorize how to work problems**	- I read books about science
- **I work with other students in small groups**	- I memorize science facts
- **I explain my answers**	- I write or give an explanation for something I am studying in science
- **I work problems on my own**	- I work science problems on my own
- **I use a calculator**	- I use a computer in science lessons
- **I use a computer**	
Metric: ***1 - "never", 2 - "some lessons", 3 - "about half the lessons", 4 - "every or almost every lesson".***	***Metric:*** ***1 - "never", 2 - "a few times a year", 3 - "once or twice a month", 4 - "at least once a week".***

Source: TIMSS 2007: User Guide. Supplement 1: Survey questionnaires.

In this dissertation, I used several approaches to measuring the level of student-centered instruction, in order to minimize the possibility of an error in estimates resulting from my measurement or variable coding choices. The items measuring the level of the

student-centered instructional climate in a classroom are examined in various ways: 1) as individual survey items, for math and science separately, 2) as composite continuous scales created based on factor analysis methodology, and 3) as binary variables that act as "yes-no" indicators of whether or not student-centered instruction was offered to students in a given classroom, again constructed through two ways: a) through an intuitive, manual coding based on my reading of the literature on student-centered instruction, and b) through collapsing a continuous factor variable that combines information from items measuring a common underlying construct.

More specifically, these items are examined as stand-alone variables in the ordinary least squares (OLS) regression analyses of TIMSS data (corrected for imputation error and clustered structure of error terms) from several education systems that are used in Chapter 6 as a starting point for examining student-centered methods. In the next step, I followed a hypothesis that student-centered instruction is a latent factor that cannot be directly measured, and various classroom methods are but proxies of this underlying construct. In order to get a gauge of the latent factor, and then to measure its relationship with student achievement in the seven post-Soviet countries including Kazakhstan, I performed exploratory and confirmatory factor analyses using the same variables measuring classroom methods as in the models described above. This process allowed me to enter student-centered instruction as a continuous composite variable, allowing for some flexibility of the relationship with achievement at different points on the continuous scale. The factor analysis process is described next.

5.3.2.2. CONTINUOUS COMPOSITE VARIABLE: FACTOR ANALYSIS

Because my goal was to examine the relationship between student-centered instructional environment as a whole, and not of one specific activity, I explored the extent to which the items described in Table 5.2 reflect a larger underlying construct. The method for accomplishing this task is exploratory factor analysis (e.g. Kim & Mueller, 1978). Factor analysis examines the variance of all items, and partitions it into *principal components*, or latent underlying factors that are not directly measurable. The variance captured by factor analysis is measured in *Eigen values*, which stand roughly for units of variance equivalent to that of one full variable. An Eigen value of 1 means that an underlying construct equivalent in variance to one measured variable has been identified. The convention of principal component analysis is that to retain a factor, its Eigen value must be at least 1. The principal components algorithm also computes *factor loadings*, or the weights of each of the individual items based on their correlation with the latent underlying factor, which are then used to reconstruct the factor into a composite variable. The reliability of the factor can be evaluated by estimating the *intra-correlation* of the items, with the *Cronbach's alpha* statistic as the indicator of whether or not the items "stick together". Conventionally, a Cronbach's alpha of at least 0.5 is required to justify the reduction of items into a composite factor, and work with single variable which acts as a common denominator, rather than with individual items (Wooldridge 2002).

While the factor analysis methodology is intended to help the researcher to identify variables that seem related, it is the theoretical and practical soundness of their combination into one factor that determine the decision to combine them into one composite variable. Items as wildly unrelated in principle as the frequency of drinking coffee and the frequency of reading Shakespeare may, in some samples, fall into one principal component, which by no means will provide a valuable measure of any common

underlying construct – at least, not one that can be easily and intuitively interpreted as a behavioral pattern, for example. Therefore, the choice of variables to include in factor analysis should be driven as much by the theoretical hypothesis, as by the practical desire to expand the number of items in a factor, thereby expanding the variance and consequently, the statistics such as Eigen values and Cronbach's alphas.

I first ran factor analysis, by country, on all items available in the student survey that measure classroom activity, at the level of individual student responses, separately for math and for science. Running the factor analysis by country, rather than as a full dataset, was a choice driven by the desire to make no assumption about the similarity of the structure of the underlying latent variable between countries. The sample sizes in each country dataset allow for a robust estimation at the country level, without aggregation. I then re-ran factor analysis on items that, according to the literature on teaching and learning, are most reflective of a student-centered approach to teaching. As an illustration, factor loadings and the Eigen values from the two specifications of the latent construct *for Kazakhstan* are presented in Table 5.3 (statistics on scale reliabilities constructed for the other countries are presented in Chapter 6, Table 6.6). Cronbach's alpha is presented for the final specifications of the factor. Pairwise deletion of missing data was chosen over listwise deletion, in order to maximize the use of all available cases, and include cases even if they are missing a data point on one of the individual items. Factor loadings were used as weights in constructing a measure of the latent variable "student-centered instruction". Each of the items were standardized to a mean of 0 and standard deviation of 1, so the coefficients on this variable will constitute expected differences in math and science achievement with each standard deviation of the treatment variable.

As can be seen in Table 5.3, a composite factor variable could only be constructed from items related to science classroom activities, but not for math teaching practices. Only four of the items related to math were representative of student-centered approach in education, and with such a low number of items, scale reliability is insufficient to reduce these items to one principal component. Therefore, these items were not combined and were either used as stand-alone predictors, or recoded into a binary variable using the procedure described below.

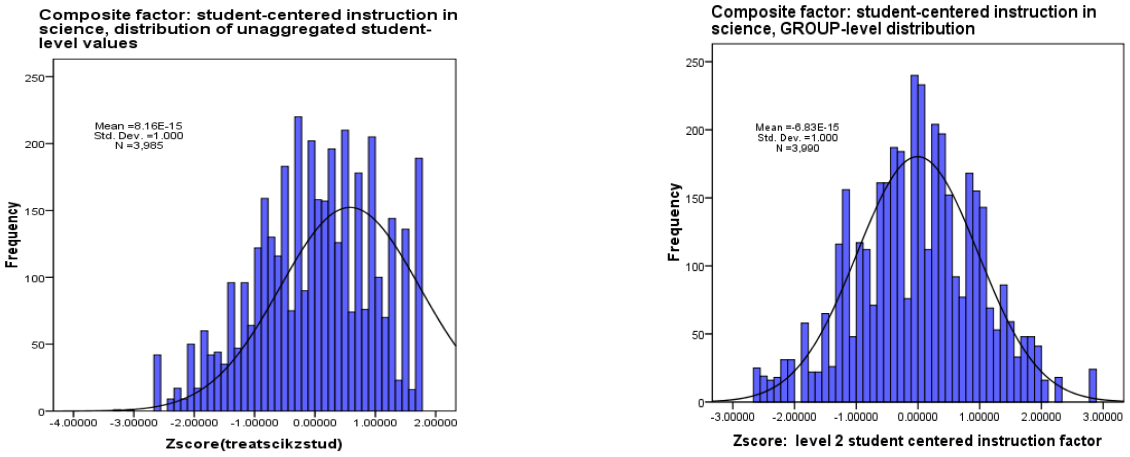
Table 5.3. Factor analysis: Student-centered instruction in math and science, KAZAKHSTAN

	Survey items	Factor loadings by item	Eigen value of first component	Proportion of variance explained by first component	Inter-item reliability (Cronbach's alpha)
Math (4 items)	I measure things in the classroom and around the school	0.41	1.6	40.1%	0.49
	I make tables, charts, or graphs	0.45			
	I work with other students in small groups	0.41			
	I explain my answers	0.30			
Science (6 items)	I look at something like the weather or plant growing and write down what I see	0.24	2.79	46.4%	0.76
	I watch the teacher do a science experiment	0.24			
	I design or plan a science experiment or investigation	0.28			
	I do a science experiment or investigation	0.28			
	I work with other students in a small group on a science experiment or investigation	0.23			
	I write or give an explanation for something I am studying in science	0.17			

The continuous factor variable is used in models that attempt to trace a linear or nonlinear relationship between the level of student-centered instruction, as a continuum stretching from "never" to "every lesson", and student achievement. While such models are generally *not causal*, with appropriate bias control they render important information that

can be followed up with an explicit causal analysis. Coefficients obtained on the continuous variable will be interpreted as being associated with each unit increase in the intensity of student-centered instruction, roughly corresponding to the possible answers to student questionnaire items outlined above in Table 5.3.

Figure 5.1. Distributions of composite variable (student-centered instruction in SCIENCE) at student and group levels.



The factor variable, estimated on individual level responses, was then pooled to the group level, using the teacher id as the grouping variable (generally, there was one teacher per class of about 25-35 students in the TIMSS sample, and one teacher per two classes in rare cases), to minimize the interference of student perceptions and obtain a more robust measure of the extent of student-centered instruction in their learning environments. For ease of interpretation, the factor variable was standardized to a mean of zero and standard deviation of 1: this scale would render the regression coefficients on the factor correspond to one standard deviation increase in intensity of student-centered instruction in science. Figure 5.1 shows the histograms of the factor variable distributions at two levels: at the student level and pooled at the group (teacher) level. It is evident that the aggregated factor

variable is more normally distributed, with groups almost equally represented on either tail of the curve.

The continuous factor variables constructed for all of the country datasets were ran as predictors of achievement in mathematics and science, as a way of looking at potential trends across the region, or identifying similarities that may be useful in further analysis of Kazakhstan's data. As I moved into a further, more detailed look at the Kazakhstan dataset, I used the continuous factor to construct a binary variable as an alternative to a more intuitive specification of "sufficient" and "insufficient" level of student-centered teaching. Through this binary coding of student-centered instruction as a treatment variable, I was able to apply a more rigorous method of bias control through propensity score matching, or the matching of cases on the background characteristics in order to strengthen the possibility of a causal inference. The process of constructing a binary variable that allowed for such a specification is described in the next section.

5.3.2.3. DISCRETE TREATMENT DEFINITION: THE BINARY VARIABLE

The binary coding of the treatment variable is based on an intuitive, rather than mechanical, combination of items, in a way that reflects "sufficient" exposure to student-centered teaching, based on the definitions found in literature (see Chapter 3). The division into the treated and control units is somewhat arbitrary, as there is no prescribed level of exposure that is guaranteed to produce results. The goal here is not to find the perfect cut between students who are considered sufficiently exposed, and those who are not, but to attempt to find a difference in math and science performance between groups of students who had had different learning experiences in their lessons, and to be able to attribute that difference to instruction. The same items as those listed in Table 5.3 were the starting point

for the coding of the binary variable. However, inclusion in the treatment group was less restrictive than would have been, had high levels of exposure been required to consider a student "treated". For the most part, students that had average class responses of above "some lessons" were included in the treatment group.

Similarly to the construction of the continuous factor, only the student responses, rather than those of their teachers, were used to create a measure of instruction in the classroom. The implicit assumption is that in the absence of an objective assessment such as third-party classroom observation, the experiences of the students with student-centered methods would be more reliably reflected by their own responses, while their teachers may communicate their intent to implement a particular method, rather than the actual implementation. The student responses, too, contain a large degree of error and noise, as 4rh graders may not always focus on the methods used in their lessons. The items themselves offer large margins of measurement error: for example, responding to the statement, "*how often / I explain what I am learning in math (science)*" may involve an in-depth discussion of the student's answers, where he or she is challenged to think critically through multiple alternatives, - or it may involve a student simply reading the proof of a theorem from their notebook.

Four items to be included in a measure of student-centered instruction were selected from the nine student survey items related to classroom activities, based on the indications from the literature of what constitutes student-centered instruction. In order to reduce error associated with individual student perceptions of what happens in the classroom, and potentially skewed depictions of the teaching environment, the responses of students were pooled at the teacher level (generally one teacher per class), so that one collective value represents the degree to which all students taught by that teacher

experience student-centered instruction. The variable measuring student-centered instruction, thus, is situated at the classroom level, while the outcomes – learning scores in math – remain at the individual level of the student.

Table 5.4. Selection of Items for the Student-Centered Instruction Index for MATH

Student Survey Items Measuring Teaching Methods	Included in Treatment Definition?	Reason for exclusion
How often I measure things in the classroom	Yes, if more often than "some lessons"	
How often I make tables, charts, or graphs	Yes, if more often than "some lessons"	
How often I work in groups	Yes, if more often than "some lessons"	
How often I explain my answers to my teacher	Yes, if more often than "about half the lessons"	
How often I work problems on my own	No	Independent activity
How often I use calculator	No	Not relevant
How often I memorize how to work problems	No	Rote memorization
How often I practice without calculator	No	Not relevant
How often I use computer	No	Not relevant
How often I work fractions and decimals	No	Content item

In this definition, students are considered "treated", or taught using some student-centered methods, if all of the conditions laid out in Table 5.4 are present: they are asked to take more often that in "some" of their math lessons, to organize information graphically more often than in some lessons, to work in groups more often than in some lessons, and to explain their answers to their teacher more often than in half of the lessons. This definition divides the samples of students in just two groups, which allows for a clean comparison of results, once the background characteristics were matched. Arguably, such a division of students into two groups based on their level of exposure to student-centered instruction is arbitrary, and there hardly exists a truly "sufficient" level of engagement in such teaching approaches to draw a line between traditional and student-centered pedagogy. However,

drawing a line at some logically justifiable point is useful for an understanding of whether there may exist gains in student achievement that can only be reached with sufficiently high intensity of student-centered pedagogy in math and science classrooms in Kazakhstan.

Table 5.5. Selection of Items for Student-Centered Instruction Index for Science

Student Survey Items Measuring Teaching Methods	Included in Treatment Definition?	Reason for exclusion
- **I look at something like the weather or plant growing and write down what I see**	In factor, but not in binary treatment	
- **I watch the teacher do a science experiment**	In factor, but not in binary treatment	
- **I design or plan a science experiment or investigation**	In factor, but not in binary treatment	
- **I do a science experiment or investigation**	Yes, if at least "once or twice a month"	
- **I work with other students in a small group on a science experiment or investigation**	Yes, if at least "once or twice a month"	
- **I read books about science**	No	Not relevant
- **I memorize science facts**	No	Independent activity
- **I write or give an explanation for something I am studying in science**	Yes, if at least once a week	
- **I work science problems on my own**	No	Independent activity
- **I use a computer in science lessons**	No	Not relevant

The definition of "treatment" for mathematics and science student-centered instruction, as described above, divided the Kazakhstan TIMSS sample into groups of treated and control students for subsequent matching on the propensity score and regression analysis (see Analysis section below). In math, the original sample put 1,934 students in the treated group, and 2,037 in the control group. In science, the breakdown resulted in 1,516 students in the treated and 2,455 students in the control group. This imbalance can be expected, because the definition of what constitutes student-centered instruction does exclude students who may have been exposed to participatory methods on an occasional basis; and given the limited acceptance of these practices in the region, no exposure or occasional exposure is true for the majority of students. Next, I explain the

process of matching these constructed sample subgroups on the estimated propensity score, with the goal of strengthening the possibility of a causal inference as a conclusion of this analysis.

5.3.2.4. PROPENSITY SCORE MATCHING

Following the division of student cases into the treated and control units on the binary treatment variable, and upon examining the descriptive statistics across the two groups, I determined that the extent to which students were exposed to student-centered instruction was nonrandom, and therefore required more rigorous bias control (see Chapter 6). I therefore chose to use the propensity score matching methodology (developed by Rosenbaum & Rubin, 1983) for restricting the sample further, and restructuring it in a way that would make the treatment groups sufficiently alike on all relevant observed background factors. This methodology seeks to approximate the experimental design, where the treatment and control groups would not be statistically significantly different on all observed measures but the treatment, such that the difference in outcomes could be directly attributed to treatment exposure. The propensity scores are estimated probabilities of students to receive treatment, calculated as a function of their background parameters (e.g. their socioeconomic status, their geographic location, the characteristics of their schools and their teachers) that predict treatment assignment, in essence condensing all of the observed relevant covariates to one value. Once units are matched on this probability, the possibility of selection bias becomes less of a concern, because the background factors correlated with treatment are equally present in units who ultimately received treatment and those who did not.

In this project, close to one hundred *probit* models with clustered standard errors were fit for math and science, using all available measures of student and teacher background and school environment, as well as higher-order terms of these variables and interaction terms among them, to estimate the probability of a given student's being in a class where instruction in mathematics and science involved greater levels of student-centered instruction. The predicted values from the probit regression were then saved as propensity scores, as they represented the probabilities of treatment occurring for an observation, given the set of covariates in the model. These propensity scores were then read in and used to match treatment and control units in the PSMATCH2 package (Leuven & Sianesi, 2003) in STATA.

The main challenge in building propensity score models is achieving *balance* on all key pre-treatment characteristics, and ensuring adequate *overlap* between the distributions of the treated and control units, so that at each point of the observed probability of treatment there is a sufficient number of counterfactuals, and no extrapolation beyond the reach of the data takes place (Gelman & Hill, 2007). Outcome models that use propensity score matched data with ideal balance on key covariates may be simple comparisons of means across the treated and control groups, but insufficient balance would call for additional control on the relevant covariates, in a multiple regression framework.

While there are several methods for matching on the propensity score, *nearest neighbor matching* was chosen as the optimal method for this analysis (see Rubin 2002; Gelman and Hill 2007; Morgan and Winship 2007). This algorithm matches each treatment unit (in this case, each student that received student-centered instruction in mathematics) with a unit in control group that has a propensity score value closest to its own. This implies that the two units in comparison have the same probability of receiving treatment

based on their pre-treatment variables, or background characteristics (i.e. treatment assignment is *ignorable*). While this assumption can never be empirically confirmed in an observational study (because of the potential that unobserved variables may have influenced selection for treatment), one can test its plausibility by comparing groups on the observed pre-treatment characteristics. Because the nearest neighbor method ties the treated units to matched controls, there may be units in the control group whose propensity score values have no matches in the treated group. In such cases, unmatched control units are deleted from the analysis because a counterfactual cannot be observed for them. Similarly, one can delete treated units with no matches in the control group, thereby restricting the analysis to the area of common support. Figure **5.2** and Figure **5.3** demonstrate the probability distributions of the cases designated as "treated" (red bars) and "control" (blue bars) prior to and after propensity score matching, for math and science student-centered instruction.

In analyses of the association between the treatment and the outcomes, propensity scores are used in several ways: one, for a comparison of means across groups in the matched sample; two, in a regression-adjusted means comparison, with additional control for background factors; and three, as probability weights in a multilevel modeling context. More on the models used for estimating the association between student-centered instruction and achievement in math and science is provided below.

Figure 5.2. Probability distribution for student-centered instruction in science, before matching.

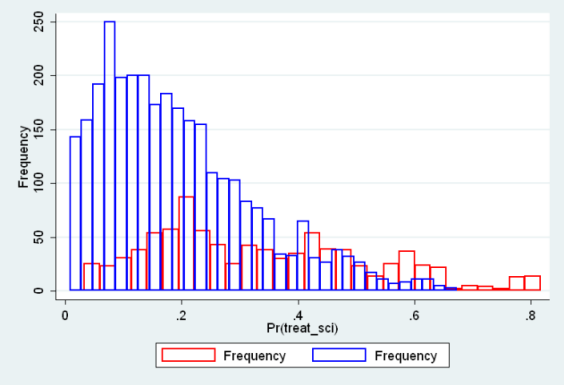
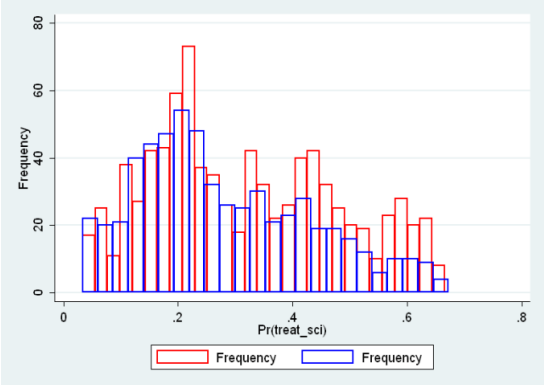


Figure 5.3. Probability distribution for student-centered instruction in science, after matching.



It is important to note once again, that the propensity score matching portion of the analysis was only conducted on Kazakhstan's data. This is one of the most labor-intensive and in-depth analytic processes, and its purpose is to go in-depth into the data and minimize all biases between the defined treatment and control groups. This procedure also allows for a better understanding of how student-centered instruction is distributed in Kazakhstan, and helps to see whether any background factors make it more likely for some students to find themselves in student-centered learning environments than others.

5.3.3. ESTIMATING THE RELATIONSHIP OF STUDENT-CENTERED INSTRUCTION WITH ACHIEVEMENT OUTCOMES

The choice of methods for the analysis of quantitative data was determined by the structure of the data and my approach to the coding of the variable of interest, student-centered instruction. The objective in employing a variety of methods was to improve the robustness of the findings, as well as exploit both the breadth and the depth of the available information.

To form an initial understanding of what types of students and teachers are more likely to engage in, or be exposed to, student-centered pedagogy, I performed simple

descriptive methods, such as an analysis of means and standard deviations on key variables. This was particularly useful prior to examining student-centered instruction as a binary variable, as I compared the treated and control students on key characteristics such as their background, their teachers' qualifications, and their school environments. This initial analysis provided important insights into the linkages between these contextual, background factors and the level of exposure to student-centered instruction, and underscored the need to control for potential selection bias when estimating the linkages between instructional methods and achievement.

In examining student-centered instruction as a continuum, I initially used ordinary least squares (OLS) regression in looking at the relationship between my variable of interest and student achievement outcomes. I first fit simple models, regressing math and science outcomes on student-centered instructional methods as individual predictors, taking items directly from the student questionnaire, and exploring the various linear and nonlinear (e.g. quadratic and exponential) forms of their relationship. Following these initial univariate analyses, I fit more saturated models controlling for observed covariates such as student background, teacher experience, the types of their community, and the presence of material shortages at their school. The uniformity of the TIMSS instruments allowed for cross-country comparison of OLS results as "snapshots" of education systems with student-centered methods embedded in an array of important predictors of student achievement, as well as placed Kazakhstan into a broader context by juxtaposing its initial findings with those from other country datasets.

The OLS regressions were estimated with robust (clustered) standard errors, taking into account the possible interdependence of unobserved characteristics of the students within schools. Furthermore, I followed the methodology for accounting for the sampling

structure through weights and jackknifing weight factors prescribed by the IEA for analyzing TIMSS data, allowed for correct adjustment for the sampling structure and the uncertainty resulting from the imputation of the dependent variable – the TIMSS test score. The dependent variable in the analyses of student-centered instruction impact on achievement was the test score, itself consisting of not one but rather five TIMSS *plausible values* for math and science, respectively. Plausible values are the test scores generated by the Item Response Theory - driven algorithm of TIMSS, based on student responses on a sample of test items during the actual test, as well as a set of predictor variables from student background surveys (Rutkowski, Gonzalez, Joncas & von Davier, 2010). None of the plausible values is the "true" test score for a given student, but collectively, the five plausible values provide a reasonably good estimate of the student's performance. Therefore, analyses with achievement as the outcome should, according to IEA, be performed on all five rather than one plausible value. This is accomplished through special customized software packages. In this study, I used STATA's plausible values procedure – the *PV* command– which executes the regression estimations five times on each of the plausible values, and adjusts the calculation of error variance to account for the imputation uncertainty around the true test score.

The regression coefficients are presented at three stages: one, with the variables measuring instruction only, then with covariates controlling for student background, covariates controlling for student and teacher background, and finally, models controlling for school-level variables measuring the overall resource environment, in addition to student and teacher background. This approach was used in fitting both the ordinary least squares with survey weighting and clustering of standard errors, as well as multilevel models (described below), with individual items measuring instruction and the composite factor for student-centered instruction (see above). While an extensive number of models

were fit on the data, only the final selection of relevant model coefficients is presented to the reader in this dissertation.

Once the initial exploratory analysis across several countries was completed, and I obtained an idea of how the different variables measuring teaching practices performed across the entire group of post-Soviet states, I proceeded to fit models on the Kazakhstan TIMSS dataset. I began with a deeper look at what, if any, background factors predict the exposure to various student-centered instructional methods and activities. After forming an understanding of what items were more or less correlated with student-centered instruction, I went on to estimate the association between student-centered instructional methods and achievement. I first ensured that selection biases were neutralized through matching to the greatest extent possible (see propensity score matching above) and then fit outcome-level models using the matched datasets.

The sample structure of the TIMSS dataset – where students were clustered into classes, which were then clustered into schools – is the classic example used in arguments for multilevel modeling in the literature, and multilevel designs such as HLM have become the norm for treating education data. Hierarchical linear modeling (HLM) is a statistical package based on a method for analyzing multilevel data developed by Steven Raudenbush and Anthony Bryk (Raudenbush & Bryk 2002). Multilevel modeling relaxes the assumption of independence of observations and, consequently, of error terms, and minimizes bias resulting from unequal group sizes, using all available information across the data and allowing both the intercept and slopes to vary across groups. Multilevel designs merge group-level models with individual-level models, which allows for maintaining the richness of individual-level variation, as well as the ability to "step back" and account for group-level coefficients without an implicit bias favoring large groups. While several software packages

exist for multilevel modeling, HLM possesses some benefits that rendered it preferable for this analysis: the possibility of running analyses on all TIMSS plausible values for mathematics and science, rather than one score at a time; the ability to accommodate survey weights, and a more explicit built-in module for modeling student-level slope coefficients as outcomes. This methodology was used in estimating the coefficients of both the continuous and the binary treatment variable in the Kazakhstan's sample, as well as the influence of student-centered instruction on the relationship between other key predictors of achievement and the achievement scores in math and science.

The HLM models were fit at two levels: the level of the student, which is where the outcome variables were located, and the level of the class/teacher, where the treatment variable was aggregated. Because in most schools, only one teacher was sampled with the TIMSS class, the school and teacher variables were collapsed into one dataset and treated as one level of data. The treatment variables – the binary specifications of student-centered instructional exposure – were entered uncentered, and the student covariates were centered at the group level. HLM models were fit on restricted (matched) samples using for the treatment variable as a binary treatment variable with two states – the treated state and the control state, on data previously matched on the propensity score using the PSMATCH2 procedure: the frequency weights calculated by PSMATCH2 were simply read into HLM as Level-1 weights for the regression model. This way, the observations within the data were restructured to ensure that each treated case had a matching control case, and unmatched control cases – or, put simply, students that could never have obtained student-centered instruction in the first place – were excluded from the comparison.

5.3.4. LIMITATIONS

The methods chosen both for the definition of the treatment variable – student-centered instruction, and for the analysis of its association with student achievement in math and science in Kazakhstan, proved to be ones that offer the most depth and minimize bias in observed data, making it possible to draw a critical assessment of student-centered approaches across a representative sample of the Kazakhstan's fourth grade population. However, important caveats must be made when interpreting the results of these analyses.

Most importantly, the nonrandom distribution of students across schools can be minimized using all available measures and statistical methods at hand, but it can hardly be completely eliminated. Even with the most rigorous bias control, such as with the propensity score matching method, treatment and control groups are not completely identical on all covariates after matching is performed, as some differences in means and standard deviations remain. In the event that there exist unobserved variables that have a strong influence on the selection of students for exposure to student-centered instruction, and the relationship between student-centered instruction and student achievement, the parameter estimates derived through regression models controlling for observed covariates may remain biased. For the validity of results presented in this dissertation, however, a lot of the variables known to exert influence on type of instruction received by a particular student, and his/her achievement outcomes, have been measured through proxies and are controlled. These proxies include student home possessions, their language of instruction, nativity to the country, the number of books in their house (intended to serve as a proxy for family cultural capital and level of education of the parents), the type of their community, size of the school, etc. However, to the extent that these proxies leave out unobserved variation – such as, for example, the level of family cultural and social capital unrelated to

the possession of books – some possibility of uncontrolled effects of extraneous factors may persist in the results.

Secondly, in the absence of pre-test measures it is impossible to isolate completely the influence of the students' individual aptitude, or their personal traits, on their achievement, and by extension, on the effectiveness of any given instructional approach in stimulating their learning. Such measures are not available in large-scale studies such as TIMSS, which are extremely costly to implement, and therefore only take a cross-section of student achievement in an education system every four years. However, the benefits of a large-scale standardized assessment tool – such as, for example, the representativeness of and unbiasedness of the sample – are substantial enough to warrant a valid exploration of the association between instruction and achievement, albeit without inferences about how student-centered instruction can improve achievement outcomes over time.

Finally, as I indicated in the beginning of this chapter, the treatment itself –namely, student-centered instructional approach – is not a homogeneous policy or package implemented in any standardized way across schools in Kazakhstan. This research project seeks to define student-centered instruction based on a mix of practices and approaches already in place in some schools, in comparison with schools or classrooms where these approaches are not as prevalent, and where traditional teacher-centered pedagogy dominates math and science lessons. The measures used to define and measure student-centered instruction are necessarily "fuzzy", with noise surrounding student reports of their classroom practices. Furthermore, there is a possibility that some of the fourth grade students misunderstood the questions in the student survey asking them about their classroom practices, and answered the items at random. Some scholars argue that the noise surrounding survey responses of younger students may be too great to be able to discern

genuine measures of reality (Levin, personal communication, 2011). To a large extent, however, the noise generated by student individual perceptions of classroom activity was reduced through pooling responses at the teacher level, so that an average response for the class serves as a proxy of the true measure of the teaching and learning environment. Nonetheless, the number of variables used to proxy the teaching environment is limited by what was captured in the student survey, which in the case of mathematics was insufficient to form a composite measure of student-centered instruction as an underlying phenomenon. A greater number of items in the student questionnaire about their learning experiences would have improved our ability to capture the latent underlying factor – but given the items currently available, inferences can only be made about more narrowly specified teaching practices.

Another caveat is that in the absence of an objective classroom observation instrument, there is virtually no information in the dataset on the quality of implementation of student centered approaches, or the attitude and perceptions of math and science teachers towards such instructional practices. This limitation is addressed to some extent by the case study on the use of student-centered instruction in schools in Kazakhstan, where I explore the thinking behind the choices (or lack thereof) of instructional methods in a math or science lesson, and the level of involvement, as well as the policy choices of the centralized state in improving the quality of math and science instruction. The methodology used for the case study is presented below, while the results are reported in Chapter 7.

5.4. DATA AND METHODS FOR THE QUALITATIVE CASE STUDY

As I explained in the beginning of this chapter, the goal of the qualitative case study of Kazakhstan's education policy and the role of its national government in education is to provide context for a deeper understanding the results of the quantitative analysis. The

interviews and documents analyzed for this part of the project sought to uncover the mechanics of state involvement in education in the recent years (2000 -2010), and evaluate the merits of a hypothesis that stronger, and more centralized, state control of the education policy and its implementation was beneficial for instructional quality, and consequently, for student achievement outcomes.

This hypothesis rests on the theoretical framework, laid out in Chapter 4, that a centralized state with a tightly coupled national education system has mechanisms at its disposal that allow it to provide a high quality of education to all its young citizens. The argument is based on the Carnoy et al. (2007) study of the Cuban education system, where the author makes a case for a highly controlled educational system with severely restricted individual choices, and a streamlined implementation of the national education policy throughout its elements, starting from curriculum development to lesson planning and classroom management. My project seeks to zoom in on the pedagogy aspect of instructional quality, and on the types of instructional methods, if any, that are endorsed and supported by the state and adopted by the teaching cadre, with specific interest in the implementation of student-centered teaching methods.

It is important to note here my own connection to the qualitative information about Kazakhstan's education system. As a native of Kazakhstan and a graduate of its education system, at the time of this data collection I already possessed an "insider's view" of the schooling practices both prior to transformation, and in the early days of independence, when the urge for a change in pedagogical practices was only beginning to be felt. Further, as a project manager overseeing donor assistance projects in education, I was acquainted with the most prevalent educational models promoted by international actors in the region, and the arguments that were made in favor of instructional changes, even though none of

the projects I oversaw worked in Kazakhstan specifically. Finally, my connection to the interviewees was made initially through teachers and administrators I knew, either through a personal or professional acquaintance.

5.4.1. QUALITATIVE DATA

The case study involved the review of key government policy documents reflecting both its education policy and the status of the education system, as reviewed and evaluated by state-supported educational analytic agencies. Documents reviewed as part of this case study were as follows:

1. The Law on Education (1998)
2. Annual Report on the State of Education 2007
3. Annual Report on the State of Education 2008
4. Annual Report on the State of Education 2009
5. National Education Development Program 2005-2010
6. National Education Development Program 2011-2020
7. Official letter from the Ministry of Education, signed by the Head of the Department of General Education

The documents chosen for review and analysis reflect the vision of the state for the development of the national education system, and presented ample opportunities to capture the priorities placed by the state on specific outcomes it seeks to achieve. Because the purpose of document analysis was the formulation of an understanding of the overall state policy in education, as well as its praised accomplishments and goals for the near future, they were used as the context for key informant and focus group interviews, rather than as sources of specific word frequency searches. The qualitative data were not

quantified in any way, which follows the recommendations in the literature on qualitative analysis (e.g. Creswell, 2008). Instead, data were used in their full scope and richness, and this information was used to answer the research questions of this study.

Table 5.6. Case study interviewees by category

Respondent category	Almaty	Astana	Bishkek
Teachers:			
- **individual**	5	3	
- **in focus groups**	5	16	
School administrators (principal & assistant principal)	5	2	
Teacher trainer (in-service)	2		
Teacher trainer (pre-service)	2		
Ministry of Education		1	
National Testing Center		1	
Student test preparation center	1		
National Center for the Quality of Education	1		
NGO officers: Kazakhstan	1		
NGO officers: Regional (Central Asia)			1
Total interviewed:			**46**

Interviewees included teachers and principals of comprehensive schools (public and private), teacher training specialists of in-service training institutes, professionals working at nonprofit institutions in education, a representative of the Ministry of Education, and an official of the National Center for Testing. Table 5.6 lists the interviewees, their category, and the type of the interview (individual or group).

Key informants for this case study were selected using purposeful sampling (Creswell, 2008). While I made reasonable effort in broadening the characteristics of the interviewees – drawing from both Russian language and Kazakh language schools, public and private schools, public agencies with a mandate in education, nonprofit groups, and private institutions offering paid education services – sampling was confined to two largest cities in Kazakhstan: Almaty and Astana, and may not, therefore, sufficiently represent the

views of the rural population, or of other ethnic minorities. These two locations provided access both to policy makers and classroom teachers, and benefited from the presence of a community of educators, and therefore, greater breadth of opinions than would be expected in smaller settings.

In recruiting interviewees, I first approached several educators known to me, and then pursued a "snowball" strategy for contacting additional interviewees, taking suggestions from those already interviewed. While all interviewees were requested 60 minutes for the interview, the actual length of interviews varied in length, ranging from 30 minutes to 1.5 hours, depending on the availability of the interviewees, the completeness of their responses, and their ability to stick to the subject of the interview. Interviews took place at locations convenient to the interviewees, in most cases their workplace, such that the setting would provide the maximum level of comfort for rich discussions.

All interviewees for the case study were informed that their responses would remain anonymous, unless they requested otherwise. Despite the relatively neutral nature of the subject in discussion, and the abundance of public debate on matters related to education, all respondents, but especially teachers, were wary of sharing their opinions with an outsider, particularly if they disagreed with existing state policy in education, were critical of the way the policies were implemented at the local level, or did not fully subscribe to the overall narrative of educational development in Kazakhstan. Almost all of the interviewed teachers were female, while the two policy makers, as well as the director of a for-profit test prep center, were male. This was not entirely unexpected, as the education sector in Kazakhstan is dominated by females, with a growing presence of males at higher levels of the career ladder. Two of the male interviewees were previously school principals, and all three had climbed the career ladder within a decade of entering the profession.

Teachers, on the other hand, were nearly always female, with rare exceptions (one male teacher participated in a focus group interview in Astana). Indeed, the TIMSS data showed that 95% of the sampled students in Kazakhstan had female teachers, and this proportion is equally high in other Eurasian countries that took part in the assessment – with Ukraine showing less than 1% of students taught by male teachers.

5.4.2. METHODS OF ANALYSIS

Because the qualitative case study focused on the state as the unit of analysis, looking at how the extent of centralization affects its ability to generate and to sustain instructional quality, the interview questions sought out the views of educators on the developments in state education policy and its implementation in the classroom. In the interviews, I solicited responses reflecting both the actual experiences with post-Soviet transformation, as well as the perceptions of educators and policy makers of the successes and challenges of education reforms in the country in the past ten years. Furthermore, as my quantitative analysis examines student-centered instruction as a relatively new approach in education, the case study interviews also sought the views of the respondents on this type of pedagogy in general, based on their knowledge, understanding, and experience with child-centered pedagogy, particularly in the areas of math and science.

Interview questions were tailored to the type of respondent: with teachers, I focused mainly on their classroom practices, their perceptions of what effective pedagogy means, whether or not they believed the instructional approach had to be changed, and their assessment of the government capacity and effectiveness of involvement in improving the quality of education. Through probing questions on the choices teachers made for their instructional approaches, I gauged their familiarity with the underlying theory behind

student-centered instruction, their personal level of comfort with such pedagogy, and their perceptions of its advantages and disadvantages of student-centered with traditional direct instruction.

In interviews with government officials (Ministry of Education, National Center for Testing, National Center for Quality Improvement, Almaty City in-service training institute), I sought out their views on government vision and policy in education, areas needing further improvement, and the opportunities and obstacles in improving classroom instruction, and consequently, student achievement outcomes. Through these questions, I sought to gauge the extent of their agency's engagement with schools, and specifically, classroom teachers; the amount of guidance and support provided to teachers, and the type of support – material or capacity development – that was seen as filling a priority gap in improving classroom instruction. In line with my theoretical framework and research questions, I looked for indications of how the level of centralization in the education system helped or hindered the development of the education system in Kazakhstan, especially when it comes to the quality of achievement outcomes, as measured by standardized assessments. The vision and priorities outlined in the policy documents I had reviewed prior to the interviews served as the starting point for building a deeper understanding through the interviews of how the state translated its vision for the development of the system into actual implementation.

Both the key informant interviews and the focus groups were conducted in a semi-structured manner. Using my interview guides as a general framework, I allowed sufficient space in the interviews and focus group discussions for the respondents to offer insights on what they felt were important developments in the public education system. For example, a new theme that emerged in several interviews was gifted education, and school tracking

policies in general. Once such a theme appeared in an interview, I followed it in subsequent interviews, in order to gain a broader perspective on the issue and to triangulate the responses. Because the primary goals of the case study were to provide context for the interpretation of quantitative results, to enrich our understanding of the "common narrative" in education in Kazakhstan (if such existed), and the role of the state in shaping the prevailing instructional environment, it was essential to preserve some degree of openness in the interviews in order to capture the broad picture.

As I mentioned above, the qualitative information collected from the interviews, as well as through the document analysis, were not quantified, but analyzed in their full breadth and richness, using the exact words and phrases of the respondents as pieces of data. Each interview was transcribed and subsequently analyzed as a written document. Repeated themes across multiple interviews were gathered as evidence of "trends" or phenomena, and subsequently condensed to the key points, presented in Chapter 6. Exact quotes from interviews were used to illustrate specific arguments, and to serve as supportive evidence for the findings presented in the case study.

5.4.3. LIMITATIONS OF THE QUALITATIVE CASE STUDY

The composition of interview respondents and the policy documents reviewed for this case study adequately serves its goals: providing context for the interpretation of quantitative findings on the relationship between student-centered instruction in math and science and achievement outcomes, and deepening our understanding of the role of the central state in shaping the instructional environment in Kazakhstan. It does not offer a comprehensive review of Kazakhstan's policy formulation and policy implementation dynamics in the education sector. Further, while the documents analyzed form the basis of

national education policy in Kazakhstan, some depth of detail of policy implementation may have escaped the analysis, as lower-level procedural documentation could not be obtained, and therefore was not reviewed.

The interviewees selected for the case study represent a diverse group of actors in the country's education sector, and for that reason, their views are likely to reflect the broad realities of Kazakhstan's education sector as a whole. However, the possibility of bias remains in favor of urban education, since no rural school teachers or principals could be accessed for an interview. At the same time, teachers in Astana, in particular, reported having experience teaching in small towns prior to their jobs at the time of the interview, and claimed to have gained a perspective on the comparison between rural and urban school settings. In addition to the heavy urban presence in the group of respondents, respondents in Almaty were slightly in the older age group than those in Astana, although none were comfortable giving their exact age.

Finally, because of the impracticality of drawing a random sample for interviews, some degree of convergence in opinions could be explained by the sampling strategy (snowball), rather than be reflective of the true set of opinions across all educators of all categories. This, however, is a risk inherent in all qualitative studies that choose purposeful sampling as a strategy, as well as in studies where the researcher has limited control over access to potential key informants.

5.5. EFFECTIVENESS OF MIXED METHODS APPROACH

Overall, the strategy I selected was effective in addressing the research questions of this study. The quantitative part of my dissertation – the analysis of TIMSS data – allowed for maximum possible objectivity in assessing the relationships between the type of

instructional methods and the student achievement outcomes. With the quantitative methods I used, I was able to capture both substantial breadth, initially examining several countries' data in addition to Kazakhstan, and significant depth, working with the Kazakhstan data to identify and minimize inherent selection biases prior to evaluating the relationship between student-centered pedagogy and achievement. While the limitations of the data itself, as I noted above, put boundaries on my ability to provide a completely conclusive answer to the question, "Is student-centered instruction effective in raising student achievement outcomes?", the methods of analysis I applied to the available data offer a substantial level of confidence in my finding of a negligible coefficient on this type of instruction, in the form that it is implemented in Kazakhstan, on student achievement outcomes in math and science.

The qualitative data, on the other hand, allowed me to examine the policy and organizational context of the education sector in Kazakhstan, as well as the perceptions, beliefs, and understandings of pedagogy that guide the teachers' choices of instructional methods. The qualitative information helped situate the quantitative findings in the broad framework of the country's developing education sector, and demonstrate the potential causes of a lack of definitive positive impact of student-centered pedagogy on achievement outcomes, and the seemingly invariable effectiveness of traditional instructional methods.

Finally, the combination of two methods of analysis, at two levels of the system, helped generate additional questions requiring subsequent further research, both in terms of the effective pedagogy, and the state-driven mechanisms that help produce the highest achievement outcomes, forming a substantial future research agenda.

CHAPTER 6.
QUANTITATIVE ANALYSIS OF THE RELATIONSHIP BETWEEN STUDENT-CENTERED INSTRUCTION AND ACHIEVEMENT

6.1. INTRODUCTION

In this chapter, I present the results of the quantitative analysis of the association between student-centered instruction in fourth grade mathematics and science and student achievement, using TIMSS 2007 data. At the core of this analysis, I explored whether the student experiences and activities in their math and science lessons, and particularly, the intensity of their exposure to active learning instructional techniques, could be responsible for some of their learning in these cognitive domains.

As is explained in Chapter 5, this section of the dissertation was intended to capture the relationship between student-centered instruction and student outcomes in the education system as a whole, rather than in a specific, narrowly defined setting, as is often the case in studies examining the effects of various teaching methods. While the "treatment" in this case is not a specific policy or a standardized package of activities, this analysis measures how student-centered methods are different in terms of their outcomes from traditional methods, if at all. It is therefore more likely than a case study, in my opinion, to provide a realistic prediction for the effects of a large-scale dissemination of student-centered methods of instruction (albeit, the way they were implemented at the time of the study), with all of its inevitable variability in the quality of implementation, its intensity, and the presence of various other confounding factors. At the same time, one cannot rule out the possibility that had there been a coherent, well-defined, standardized set of student-centered practices, and a uniform quality of its implementation across classrooms, the results of this analysis would have been entirely different. The potential of

just such a coherent set of practices appearing in Kazakhstan, the focus country of this dissertation, is being evaluated through qualitative methods in Chapter 7. This chapter focuses on classroom practices present at the time of the assessment in 2007, and using the available data.

Three challenges inherent in the analysis of large-scale survey and achievement data determined the selection of methods for this study. First, the level of student-centered instruction was not observed directly in the TIMSS sampled classrooms, and therefore no ready measures of "student-centeredness" or, on the contrary, of "traditionalism" in math and science teaching existed in the dataset. The measures used in my analysis were constructed by me from TIMSS background student questionnaires, either in their original form or as composites of several items reflecting a broader underlying phenomenon. Chapter 5 provides extensive detail on the methods for the construction of these measures of student-centered instruction, but here it will suffice to say that the full extent of potential relationships with achievement outcomes was explored, where student-centered instruction was reflected by individual items, continuous composite factors, and dichotomous treatment variables.

Secondly, because the TIMSS data is non-experimental, a key goal of the methods used for this analysis was minimizing selection bias. It is well known that most communities, the selection or assignment of student to schools or teachers is never purely random: families with higher social status tend to send their children to better resourced schools, and may make informed choices about the types of instructional methods to which their children are exposed. Such families are also more capable of creating an environment in their homes that contributes positively to their children's learning progress, making it difficult to discern the home and family influences from the effects that are caused by the

school system – which are, arguably, of substantially greater interest to educators and policy makers. While completely eliminating selection bias inherent in observational studies is hardly ever possible, one can minimize the interference of such bias in estimating the associations with other variables, through stricter controls and modeling methods that approximate the experimental design. Such methods are used in this dissertation.

Finally, the hierarchical, nested structure of the data must be taken into account. In TIMSS, schools were randomly sampled from a defined population of schools in each country, and then intact fourth grade classrooms – one or two per school – were sampled within each country. Statistical analyses of the TIMSS samples that treat students as independent random samples are bound to underestimate the measurement noise and variance, and therefore may lead to Type I error, with researchers incorrectly concluding that an effect is present in the population when in fact it is not. In this dissertation, I follow the specifications of this sample to properly account for the mutual dependence of units in each group, both through the application of appropriate sampling weights designed by IEA for use with TIMSS data, and through the modeling techniques that explicitly account for the nested structure of the data.

While the main focus of this dissertation is examining the impact of such instruction on achievement specifically in Kazakhstan, in this chapter I begin with an analysis of achievement data of six post-Soviet countries in addition to Kazakhstan. These include: Armenia, Georgia, Latvia, Lithuania, Russia, and Ukraine. I believe that it is useful to first examine the parameters of the variables of interest in context, across several countries that have similar recent history to Kazakhstan and have to a large extent inherited similarly structured education systems. I offer a fairly extensive overview of the observed levels of instructional intensity in student-centered methods, and its correlation with (or lack

thereof) student achievement outcomes, as measured by TIMSS fourth grade math and science tests, and demonstrate the variability of coefficients on student-centered methods across countries. I then follow the initial findings with further analysis of the Kazakhstan's data, applying heavier controls of selection bias, such as propensity score matching, and more extensive control of the sampling structure, through a multilevel modeling approach. All of these methods are intended to isolate the effects of external variables, both on their likelihood of being exposed to such methods as opposed to traditional instruction, and on their overall achievement in mathematics and science.

As this chapter will demonstrate, I find no statistically significant relationship between student-centered methods and achievement outcomes in math and science in Kazakhstan, while the traditional methods of instruction, such as independent work on math and science problems or memorization have shown a fairly robust association, controlling for student background and appropriate school and teacher characteristics. There are several potential explanations for this lack of a definitive causal effect of student-centered methods, and I offer some of them at the end of this chapter. The following chapter explores the broader context of the education sector in Kazakhstan, including the perceptions of effective instruction on the part of the teachers of math and science, and the priorities, programs, and policies of the state as a key player in education. These qualitative findings may also be useful in understanding the results of quantitative analysis.

6.2. STUDENT-CENTERED INSTRUCTION IN A REGIONAL PERSPECTIVE

6.2.1. THE TIMSS DATA AND VARIABLES

This part of the study examines student-centered instructional methods in math and science across the seven post-Soviet countries, juxtaposing Kazakhstan against its Eurasian neighbors Armenia, Georgia, Russia, Ukraine, Latvia and Lithuania. These countries inherited largely the same education system after the breakup of the Soviet Union in 1991, and while they have for the past twenty years been implementing reforms, substantial structural similarities still remain. As Chapter 2 notes, traditional teacher-centered instruction was prevalent in the Soviet period, and the era of reform in education in these countries opened the doors to educational innovation and child-centered approaches in teaching, both in primary and secondary school. In this analysis, I explore whether there is any preliminary evidence of the impact of child-centered instructional environments on math and science test scores of fourth graders, if so, what is the level of consistency or variation in the magnitude of this impact across the seven countries.

Table 6.1. Descriptive statistics on key outcome variables by country.

	Math		Science		
	Mean score	**Standard deviation**	**Mean score**	**Standard deviation**	**Sample size**
Armenia	499.51	89.53	484.39	118.80	4,079
Georgia	438.40	88.42	417.60	84.68	4,108
Kazakhstan	549.35	83.82	532.83	74.34	3,990
Latvia	537.87	71.66	542.60	66.64	3,908
Lithuania	529.80	75.77	514.21	65.20	3,980
Russia	544.05	83.38	546.23	80.53	4,464
Ukraine	469.00	84.49	417.60	84.68	4,292

Note: Cell entries are means and standard deviations adjusted for sampling weights.

Because the basic descriptive statistics of the seven country samples were provided in Chapter 5 (Table 5.1), I begin by providing the basic information about the performance of the seven Eurasian countries on TIMSS, and their use of student-centered instruction in math and science lessons. Table 6.1 shows the means and standard deviations of the TIMSS math and science scores by country. As the table demonstrates, Kazakhstan is among the top performers both in mathematics and in science fourth grade assessments, along with Russia, Latvia, and Lithuania. The remaining three countries are spread further apart, Armenia scoring close to the international mean of 500, albeit with much higher variance in science than in math, and Georgia and Ukraine showing substantively lower scores in both subjects, well below the international mean.

The variance of individual scores around the mean also differed in this sample of countries, ranging from approximately 65 points in Lithuania's science assessment, to close to 120 in the case of Armenia. If gauged against the international standard deviation of 100 points, this spread in variances is substantial. In assessments of achievement outcomes, one wishes to find a high mean score combined with small variance, and it seems that it is best achieved in the case of Latvia, both in mathematics and science. However, I will not focus here on the direct evaluation of each country's performance on TIMSS, and on all of the potential factors that contributed to it. The sole interest of this chapter is the association of student-centered and traditional instruction and achievement. These statistics on performance on TIMSS are presented here as a frame of reference for interpreting the coefficients on the variables measuring student-centered instruction on achievement, which will be presented in regression analysis. Real coefficient sizes are often rescaled as proportional to the standard deviation of the dependent variable, so these statistics are useful for that purpose.

The variables I use as "treatment", or ones that I refer to as reflective of student centered instruction, are taken from the student background questionnaires administered alongside the student assessment in TIMSS, and are, consequently, student reports of activities they engaged in during their math and science lessons. These student reports are aggregated to the teacher level to minimize noise associated with individual perceptions and experiences in the classroom, and then applied in regression analysis as predictors of achievement. In cross country analysis, my goal is to take a "snapshot" that puts Kazakhstan alongside other countries for a comparison of observed coefficients, and therefore, only two of the coding methods for treatment variables are used – one, with items as separate predictors controlling for each other, and two, in science instruction, as a single composite factor reflective of student-centered instruction as an underlying construct. The detail on the selection of the variables from the student questionnaires is provided in Chapter 5.

It is useful as a first step to examine the mean level of exposure to the teaching practices of interest across these countries. If one accepts the hypothesis of a common starting point for these countries – the dissolution of the Soviet education system – the results that we see in the teaching practices may reflect the differences that have taken place in the past twenty years. This hypothesis can hardly be tested or proven right, however – the differences in teaching practices, if not in institutions and structures, may have been present long before these countries gained independence. Regardless of the timeframe in which the divergence of practices took place, a first look at the mean frequencies on each of the teaching practices, presented in Table 6.2 reveals a seeming consistency: the standard deviation of the country-level mean values on all of the methods in the table is 0.2, on a scale from 1 to 4, and values generally fall within the range of 2 ("some lessons") and 3 ("about half the lessons"). However, multiple comparisons tests on

these mean values (Bonferroni tests) showed that in most cases, countries are statistically significantly different from each other in the frequencies of exposure to the different teaching practices, as reported by their students.

Table 6.2. Country means of the key teaching methods in math and science lessons, as reported by students

Subject	Questionnaire item	Method	ARM	GEO	KAZ	LAT	LIT	RUS	UKR
MATH									
Metric: 1 - "never", 2 - "some lessons", 3 - "about half the lessons", 4 - "every or almost every lesson".	I measure things in the classroom and around the school	Measurement	1.77	2.13	2.13	1.86	1.79	1.88	2.35
	I make tables, charts, or graphs	Graphic organization	2.36	2.52	2.47	2.12	2.64	2.42	2.28
	I work with other students in small groups	Group work	2.38	2.64	2.53	2.25	2.35	2.10	2.47
	I explain my answers	Reflection and feedback	3.22	3.31	3.42	2.84	2.90	3.26	3.35
	I memorize how to work problems	Memorization	3.32	3.52	3.58	3.53	3.45	3.35	3.42
	I work problems on my own	Unassisted independent work	3.34	3.53	3.61	3.08	3.54	3.06	3.21
SCIENCE									
Metric: 1 - "never", 2 – "a few times a year", 3 - "once or twice a month", 4 - "at least once a week".	I look at something like the weather or plant growing and write down what I see	Observation 1	2.66	2.71	3.01	2.87	2.70	2.68	2.97
	I watch the teacher do a science experiment	Observation 2	2.92	2.63	3.00	3.06	3.17	2.71	2.94
	I design or plan a science experiment or investigation	Project-based learning 1	2.20	2.10	2.61	2.76	2.19	2.22	2.50
	I do a science experiment or investigation	Project-based learning 2	2.07	1.99	2.58	3.02	2.14	2.16	2.49
	I work with other students in a small group on a science experiment or investigation	Group work	2.12	2.11	2.62	2.72	2.31	2.04	2.41
	I write or give an explanation for something I am studying in science	Reflection and feedback	3.09	3.21	3.35	2.81	3.16	3.44	3.32
	I memorize science facts	Memorization	3.14	3.48	3.34	3.08	3.33	3.40	2.95
	I work science problems on my own	Unassisted independent work	2.96	3.47	3.61	3.68	3.59	3.34	3.46

The next step for a cross-country analysis is to examine whether these differences in teaching and learning environments are correlated with achievement outcomes measured by math and science tests.

6.2.2. STUDENT-CENTERED LEARNING IN A CROSS-COUNTRY REGRESSION FRAMEWORK

As outlined in Chapter 5, I use several methods to examine the association between the teaching approaches in math and science, and the achievement outcomes. I begin with the classic ordinary least squares regression, but follow the methodology recommended by IEA and TIMSS/PIRLS International Center in applying the appropriate weights and jackknifing procedures that adjust the dataset to represent the population of fourth graders in each country. The jackknifing and weight variables also correct the estimation of error variance due to imputation inherent in the test scores: each test score is not a true score for a particular student, but is a set of plausible values randomly drawn from a distribution of values generated using the multiple imputation method from the small sample of items administered during the test and a set of core background variables from the student questionnaire (Rutkowski et al. 2010). Throughout the analyses presented in this chapter, the models were run using the special "plausible values" procedure in STATA, or replicated across all plausible values and averaged to one point estimate and standard error estimate.

Table 6.4 and Table 6.5 display the coefficients on the teaching methods in math (Table 6.4) science (Table 6.5), in a regression where the outcome is the TIMSS score on the mathematics and science assessments, respectively. The regressions were run using the method described above, with an OLS model adjusted for survey weights and with clustering-robust standard errors, in three stages:

- first, with the teaching methods as the only predictors of achievement outcomes;
- second, with covariate adjustment for student background variables (age, gender, their nativity to the country, all available measures of household income, and their emotional experiences at school, i.e. being bullied) and basic school characteristics, such as the size of the community in which it is located, and the total enrollment of the school;
- and third, with covariate adjustment for the student and school variables, plus adjustment for basic teacher variables, such as teacher age (with a square term), their gender, level of education, length of teaching experience (also with a square term), their certification status, and the intensity of their interactions with peer teachers at their school.

Table 6.3 shows the variables used as covariates in the OLS models examining the relationship between student-centered methods of instruction and achievement outcomes.

Table 6.3. Control variables used in ordinary least squares regressions.

Variable Name	Description	Type
AS4GSEX	Gender of student	Binary
ASDAGE	Age of student	Continuous
AS4GBORN	Student born in country	Binary
AS4GBOOK	Number of books at home	Ordinal, 5 categories (1="none", 5="three or more bkcases")
revAS4GSTOL	Student had something stolen at school	Binary
revAS4GHURT	Student was hurt at school	Binary
revAS4GMADE	Student was made to do things at school	Binary
revAS4GMFUN	Student was made fun of at school	Binary
revAS4GLEFT	Was left alone by classmates	Binary
AS4GTH01	Student has calculator at home	Binary
AS4GTH02	Student has computer at home	Binary
AS4GTH03	Student has study desk at home	Binary
AS4GTH04	Student has dictionary at home	Binary
AC4GTENR	Grade 4 enrollment at school	Continuous
_IAC4GCOMU_2	city of 100K- 500K residents	Binary
_IAC4GCOMU_3	city of 50K to 100K residents	Binary
_IAC4GCOMU_4	town of 15K to 50K residents	Binary
_IAC4GCOMU_5	town of 3K to 15K residents	Binary
_IAC4GCOMU_6	village less than 3K residents	Binary
AT4GAGE	Teacher age	Ordinal, 6 categories (1="Under 25", 6="60 or older")
AT4GTAUT	Teacher experience in years	Continuous
AT4GSEX	Teacher gender	Binary

Table 6.3., continued

AT4GTLCE	Teacher certification status	Binary
_IAT4GFEDC_3	Teacher had finished ISCED 4 equivalent education	Binary
_IAT4GFEDC_4	Teacher had finished ISCED 5B equivalent education	Binary
_IAT4GFEDC_5	Teacher had finished ISCED 5A equivalent education	Binary
_IAT4GFEDC_6	Teacher had finished ISCED 5A second degree equivalent education	Binary
AT4GTAUTsqrd	Teacher experience, squared	Continuous
AT4GAGEsqrd	Teacher age, squared	Continuous
AT4GOTDC	Frequency of interaction between teachers on concepts	Ordinal, 4 point (1="Never", 4="Daily or Almost daily")
AT4GOTPM	Frequency of interaction between teachers on preparing for lessons	Ordinal, 4 point (1="Never", 4="Daily or Almost daily")
AT4GOTVT	Frequency of interaction between teachers, visits to classroom	Ordinal, 4 point (1="Never", 4="Daily or Almost daily")
AT4GOTAT	Frequency of interaction between teachers, informal observation	Ordinal, 4 point (1="Never", 4="Daily or Almost daily")

By gradually adding covariates to the model, I examine the sensitivity of coefficients to the model specifications. If the coefficients do not change, one may conclude that the covariates are not interfering with the association (or lack thereof) of the teaching practices and student achievement. The results in Table 6.4 and Table 6.5 show that there is no single method that is associated with an overwhelmingly large positive or negative differences in test scores, be in mathematics or in science. In mathematics, the more student-centered, active learning methods, such as taking measurements during class, or graphically organizing information into tables, charts, and graphs, are overwhelmingly showing either a null association with test scores, or in the case of Kazakhstan, a strong and persistent negative association. Somewhat counter-intuitively, the interpretation of the coefficient on "measuring things in the classroom and around the school" is that with each step upwards, on the scale of 1 to 4, in the level of frequency of this activity, we notice a drop in mean mathematics scores by a dramatic 30-38 score points, which is roughly 35% of a standard deviation for Kazakhstan. This negative association persists when controlled for student, school, and teacher background variables. This is the largest negative coefficient on this activity across the seven countries represented in the sample, although all point estimates

are negative; only Ukraine exhibits a negative coefficient that passes the conventional level of significance testing, but with a point estimate roughly a third of Kazakhstan's.

Graphic organization in the math lessons – "making tables, charts, graphs" – has shown no effect in most countries, including Kazakhstan, but was associated with an additional 24 score points in Russia, or about 1/3 of its standard deviation. This means that going from "never" having to make tables, charts, or graphs in a math class to doing it in "about half the lessons" shows almost a 50-point bump in the math score, which quite a substantial magnitude of a difference. Again, these coefficients have to be interpreted with caution as *observed differences in mean scores* by intensity of exposure to the teaching activities, at the class-level, and not as causal effects.

Another classroom activity variable that had a large point estimate for the slope coefficient in Kazakhstan was "working on math problems on my own": ranging from 44 to 33 test score points, depending on level of covariate adjustment. However, the standard errors for this estimate were also very large, making it impossible to reject the null hypothesis when this variable is controlled for student and teacher background. In other countries, working on math problems independently produced different results: large, positive, and significant in Lithuania; and insignificant, albeit with positive point estimates, everywhere else in this cross-country sample. Similarly, "memorizing how to work math problems" did not offer a conclusive result, positive or negative, ranging from large and negative significant (Russia) to large positive significant (Lithuania), and insignificant in all other countries, with highly varied coefficient values.

Table 6.4. Regression coefficients on MATH methods of instruction, with varied levels of adjustment for covariates.

Level of covariate adjustment	Classroom Activity	Armenia		Georgia		Kazakhstan		Latvia		Lithuania		Russia		Ukraine	
		coeff	se	coeff	se	coeff	se	coeff	se	coeff	se	coeff	se	coeff	se
Unadjusted	Measurement in the classroom	-18.81	(10.03)	-12.98	(10.43)	*-37.29*	*(13.12)*	-10.70	(7.36)	4.13	(7.46)	-10.98	(11.84)	-20.89	(10.23)
	Making tables, charts, and graphs	5.62	(10.57)	-4.24	(12.32)	1.97	(11.27)	13.22	(8.91)	-2.23	(7.02)	28.32	(6.54)	-4.58	(11.19)
	Working in small groups	11.92	(11.67)	3.60	(8.18)	4.04	(15.02)	-10.21	(7.42)	*-32.24*	*(5.79)*	-18.43	(12.53)	-10.39	(8.14)
	Explaining back to the teacher	-11.26	(15.22)	13.09	(15.48)	18.26	(19.39)	-17.88	(9.21)	8.00	(6.85)	10,41	(9.29)	24.08	(12.40)
	Working on math problems on your own	8.24	(13.22)	-7.37	(16.00)	*42.65*	*(21.78)*	*39.06*	*(9.90)*	*42.78*	*(9.88)*	30.58	(14.42)	30.88	(12.20)
	Memorize how to work problems	-18.33	(17.28)	21.93	(18.64)	-21.06	(22.15)	13.61	(6.41)	*28.41*	*(10.56)*	-25.80	(7.03)	5.70	(6.42)
Adjusted for student background, community size, and student social comfort at school	Measurement in the classroom	-21.88	(10.61)	-19.18	(9.28)	*-38.56*	*(11.37)*	0.81	(7.67)	3.87	(5.55)	-8.65	(13.29)	-17.86	(8.01)
	Making tables, charts, and graphs	16.53	(11.74)	-1.54	(10.63)	6.10	(12.84)	3.37	(7.80)	-9.14	(6.42)	23.52	(6.37)	2.68	(7.34)
	Working in small groups	16.47	(8.75)	06.18	(8.00)	-0.25	(15.79)	-4.04	(6.64)	*-13.00*	*(4.90)*	-9.05	(12.13)	0.34	(6.31)
	Explaining back to the teacher	-5.18	(14.58)	-4.44	(13.00)	15.54	(16.55)	-5.51	(8.27)	1.31	(6.13)	10.77	(9.95)	*20.60*	*(11.95)*
	Working on math problems on your own	16.80	(14.51)	7.83	(15.86)	34.49	(20.39)	*19.73*	*(8.38)*	*25.66*	*(7.01)*	10.17	(15.26)	9.53	(9.12)
	Memorize how to work problems	-22.21	(20.88)	3.68	(16.41)	-25.69	(21.57)	6.65	(5.12)	*22.07*	*(9.73)*	*-26.62*	*(6.22)*	-0.55	(5.33)
Adjusted for student background, community size, student social comfort at school, and teacher characteristics	Measurement in the classroom	-11.84	(11.28)	-6.59	(11.16)	*-28.73*	*(13.77)*	5.45	(6.74)	5.77	(6.17)	-8.65	(13.55)	-17.25	(8.66)
	Making tables, charts, and graphs	12.39	(11.75)	-1.68	(12.00)	-7.28	(12.07)	2.18	(7.22)	-9.94	(6.80)	24.28	(6.66)	-0.19	(8.23)
	Working in small groups	13.33	(10.24)	3.84	(8.42)	9.46	(14.86)	-4.19	(7.02)	-16.55	(6.40)	-3.21	(13.96)	01.10	(6.48)
	Explaining back to the teacher	-13.20	(15.35)	-0.61	(14.21)	4.16	(14.21)	0.14	(9.44)	4.87	(7.11)	13.36	(10.73)	22.02	(12.23)
	Working on math problems on your own	16.49	(15.32)	02.51	(18.22)	32.96	(18.21)	18.28	(7.74)	*30.81*	*(7.66)*	13.43	(15.23)	8.83	(8.80)
	Memorize how to work problems	-22.34	(18.84)	-0.08	(18.55)	-9.61	(18.27)	6.36	(5.41)	*23.07*	*(9.36)*	*-27.93*	*(7.33)*	-0.21	(5.35)

Note: Dependent variable: TIMSS Mathematics test score (5 plausible values). Cell entries are regression coefficients, with standard errors in parantheses. Italicized values are statistically significant at p<.05 level.

Table 6.5. Regression coefficients on SCIENCE methods of instruction, with varied levels of adjustment for covariates.

Level of covariate adjustment	Classroom Activity	Armenia		Georgia		Kazakhstan		Latvia		Lithuania		Russia		Ukraine	
		coeff	se	coeff	se	coeff	se	coeff	se	coeff	se	coeff	se	coeff	se
Unadjusted	Observation	-16.82	(15.99)	-19.33	(9.08)	*-3.10*	*(10.21)*	-17.58	(9.50)	-14.93	(6.42)	-4.43	(9.73)	-1.60	(12.43)
	Watch experiment by teacher	-26.65	(17.04)	10.96	(8.98)	-16.45	(10.34)	9.55	(6.68)	-6.00	(5.00)	-12.96	(10.14)	4.58	(11.28)
	Plan experiment	-4.17	(20.29)	-14.12	(16.36)	1.87	(24.50)	2.00	(9.18)	*-6.57*	*(7.81)*	-20.64	(13.33)	-6.67	(14.27)
	Do experiment	27.69	(22.06)	24.83	(18.42)	-27.81	(25.88)	-5.54	(8.28)	1.41	(6.38)	32.07	(10.59)	4.44	(12.62)
	Work in small group	-12.80	(17.43)	-23.35	(9.82)	*21.44*	*(9.76)*	*1.51*	*(5.70)*	*0.13*	*(4.81)*	3.55	(7.64)	-13.88	(9.82)
	Give explanation of what I'm studying	18.26	(16.13)	-9.53	(13.87)	-10.35	(15.00)	-23.74	(7.12)	*-6.50*	*(6.25)*	25.53	(13.76)	-16.48	(17.54)
	Memorize sci facts	16.65	(18.24)	2.66	(15.41)	*29.04*	*(12.47)*	21.57	(6.57)	34.51	(7.14)	-32.15	(8.07)	16.07	(12.00)
	Work science problems on own	1.07	(14.99)	11.21	(14.93)	38.94	(14.68)	23.48	(11.52)	21.13	(7.27)	22.24	(9.11)	29.24	(23.33)
Adjusted for student background, community size, and student social comfort at school	Observation	-13.78	(19.92)	-22.37	(7.01)	-5.15	(10.60)	-11.18	(8.59)	-4.78	(5.34)	1.40	(8.89)	0.92	(9.93)
	Watch experiment by teacher	-16.03	(17.55)	13.54	(8.36)	*-20.66*	*(10.62)*	7.58	(5.20)	-3.30	(4.12)	-10.28	(8.85)	2.41	(6.74)
	Plan experiment	5.10	(23.98)	-13.03	(13.18)	7.29	(23.47)	2.73	(8.68)	-1.73	(5.82)	-18.75	(11.12)	-6.45	(9.01)
	Do experiment	14.40	(23.27)	10.17	(12.53)	-28.05	(23.24)	-0.50	(7.78)	-2.88	(5.10)	*20.09*	*(8.54)*	2.73	(7.76)
	Work in small group	-5.43	(17.08)	-10.11	(8.63)	*18.34*	*(9.92)*	2.76	(5.66)	-0.42	(3.42)	7.56	(7.39)	-1.89	(6.11)
	Give explanation of what I'm studying	5.20	(19.55)	-11.94	(13.75)	-8.26	(14.92)	*-13.15*	*(8.03)*	0.42	(5.08)	23.16	(14.80)	-1.10	(13.86)
	Memorize sci facts	32.50	(25.06)	-3.83	(14.73)	*27.05*	*(11.82)*	*13.48*	*(6.84)*	*20.33*	*(6.14)*	*-29.31*	*(8.01)*	7.61	(9.43)
	Work science problems on own	5.63	(16.67)	7.91	(12.78)	*37.92*	*(12.40)*	8.98	(9.67)	18.68	(7.01)	11.94	(10.18)	2.18	(19.82)
Adjusted for student background, community size, student social comfort at school, and teacher characteristics	Observation	-17.36	(19.09)	*-24.74*	*(9.20)*	-0.62	(9.82)	-10.86	(8.93)	-7.16	(6.48)	5.74	(9.33)	0.43	(10.44)
	Watch experiment by teacher	-16.50	(16.96)	9.92	(10.00)	-9.47	(11.99)	2.78	(5.44)	-2.72	(4.10)	-2.68	(8.18)	0.92	(6.81)
	Plan experiment	40.07	(24.91)	-11.90	(16.04)	-7.99	(17.99)	-0.73	(9.31)	0.14	(6.27)	-26.32	(11.30)	-5.13	(9.22)
	Do experiment	4.05	(22.65)	12.57	(13.20)	-29.69	(18.91)	0.84	(8.04)	-3.38	(5.38)	22.31	(9.64)	0.60	(9.24)
	Work in small group	-18.18	(17.33)	0.20	(11.97)	*21.66*	*(9.28)*	1.04	(5.63)	-1.26	(4.38)	2.38	(7.08)	-1.50	(6.43)
	Give explanation of what I'm studying	1.70	(19.11)	-11.46	(16.11)	0.04	(14.80)	-5.76	(8.26)	1.37	(5.18)	19.53	(17.25)	-2.75	(13.97)
	Memorize sci facts	24.57	(22.63)	0.42	(15.77)	*26.66*	*(11.64)*	11.00	(7.75)	*21.52*	*(6.69)*	-26.90	(8.38)	8.55	(8.98)
	Work science problems on own	фев.15	(16.02)	6.73	(16.64)	*31.82*	*(12.00)*	14.19	(9.93)	*26.30*	*(7.77)*	20.58	(10.09)	0.05	(19.19)

Dependent variable: TIMSS Science score (5 plausible values). Cell entries are regression coefficients, with standard errors in parantheses. Italicized values are statistically significant at p<.05.

Interestingly, in Armenia and Georgia data, the teaching methods in mathematics explain precisely nothing in the variation of student achievement outcomes on the TIMSS math test. None of the variables, whether they are classified as student-centered or traditional, showed a statistically significant coefficient in any of the regression models. On most variables in the three model specifications (teaching practices only, adjusted for student background, and adjusted for student, teacher, and school background) in Armenia and Georgia, the standard errors were equal to or larger than the point estimates, which means that the uncertainty around the association between the methods and the learning scores was too great to conclude whether such an association exists in the population.

In regressions of *science* scores, the situation is quite similar – with a lot of unstable coefficients and large standard errors on most variables. *Watching a teacher perform an experiment, planning an experiment,* or *doing an experiment* did not show statistically significant associations with science test scores anywhere except Russia, controlling for other variables and covariates of student and teacher variables. In Kazakhstan, point estimates on these variables are negative, when controlled for student and teacher background variables, which means that in this sample, students who were doing or planning more experiments were performing worse than those who were engaged in experiments less, but that correlation could not be generalized to the overall population. In Armenia, the point estimate on planning an experiment dramatically increased to roughly 40 points when controlled for all available covariates, but with a standard error firmly around 25, the null hypothesis again could not be rejected, and we conclude that there may not be any association in the population of Armenia's fourth-graders.

Group work in planning or conducting an experiment, *however, showed a very stable positive association in Kazakhstan's data, but nowhere else.* To interpret this linear

coefficient, the students in Kazakhstan who were engaged in group work about once or twice per month were likely to score, on average, 22 points higher than those who did group work only a few times a year, controlling for other teaching practices in their science lessons, as well as student background, teacher qualifications, their school size, and community. Similarly, students in Kazakhstan who were conducting group work in their science lessons at least once a week were likely to score, on average, 44 points higher than those who only did it a few times a year. In other countries, group work in conducting experiments in science lessons showed no association with science scores.

While the high point estimate for group work in Kazakhstan might be good news for supporters of student-centered instruction in this country, it must also be noted that the two methods in the category of "rote learning" – memorization of science facts and working through science problems independently – showed large and statistically significant point estimates, as well, for Kazakhstan. In fact, those coefficients are *larger than the coefficient for group work*, albeit with slightly larger standard errors, as well. These two methods for rote learning in science were also positive, large, and significant in Lithuania, and were inconclusive in the rest of the countries. For Latvia, the model with teacher background controls was fit without a control for teacher certification status, because it was not measured in that country – presumably, all teachers in Latvia are certified. A model controlled for the remaining teacher variables (age, gender, experience, level of formal education, extent of collaboration with peers) resulted in reduced point estimates and loss of statistical significance for the "rote learning" methods in Latvia. In contrast to Kazakhstan, in Russia the point estimates for memorization was negative around 27 points, with a standard error of slightly over 8, passing the conventional confidence level for statistical significance. Students who had to memorize more science facts in Russia were scoring lower, on average, than those who had to memorize less.

As a conclusion to this "snapshot analysis", I note that the regressions of math and science learning scores from the TIMSS assessments on teaching practices showed no clear patterns across countries. For the most part, the standard errors – when the sampling structure and imputation uncertainty are properly accounted for, are quite large in all countries, making it impossible to reject the null hypothesis of no association of teaching practices and achievement in the population, despite some very large point estimates. Moreover, the coefficients vary not only in their magnitude, but also in their sign across countries, with some methods showing large positive and negative associations, depending on the country. In sum, so far there is no conclusive evidence that the use of student-centered methods was associated with higher achievement outcomes in math and science in the seven post-Soviet states.

6.2.3. MEASURING STUDENT-CENTERED INSTRUCTION AS A SINGLE LATENT VARIABLE

As a next step to this cross country analysis, I hypothesize whether the variables measuring student-centered instruction are in fact measuring the same thing, and hence might be endogenous to each other. If this hypothesis were true, one would not be able to separate their respective associations with achievement, and the increased error variance would make it impossible to identify with certainty whether the relationship between these variables and the outcomes exist. The method for testing this hypothesis is: first, through scale reliability tests, examining how well the items fit together using Cronbach's alpha as a measure of inter-item reliability; and secondly, through factor analysis, where the total variance of all of the items combined is partitioned into underlying latent "factors" which are uncorrelated with each other. If at least one latent factor can be identified – which, conventionally, would be the case if the reported Eigenvalue for that factor is higher than 1

– the variables can be collapsed into one composite measure, and each item enters with a weight, estimated as part of the factor analytic procedure (see Chapter 5 for more detail).

I tested the hypothesis on both the math and the science achievement outcomes. In general, the larger the number of items in a given scale, provided they do measure one and the same construct, the higher the inter-item reliability, and the stronger the odds of detecting a robust underlying construct. With only four items pertaining to student-centered methods in mathematics lessons, the chances of finding a reliable scale or factor are low. For science methods, however, a larger number of items in the questionnaire were measuring teaching practices, to begin with, and consequently a larger number of items fell into the "student-centered activities" category. The scale reliability statistics are presented in Table 6.6.

Table 6.6. Scale reliability (inter-item correlation) statistics for math and science teaching methods

	Math		**Science**	
	All items (6 items)	**Student-centered methods (4 items)**	**All items (8 items)**	**Student-centered methods (6 items)**
Armenia	0.567	0.491	0.736	0.695
Georgia	0.571	0.528	0.737	0.744
Kazakhstan	0.549	0.496	0.755	0.764
Latvia	0.565	0.534	0.718	0.713
Lithuania	0.567	0.540	0.736	0.734
Russia	0.473	0.469	0.722	0.715
Ukraine	0.553	0.490	0.749	0.723

Note: Cell entries are Cronbach's alpha coefficients of interitem correlation at the individual student level. Cronbach alpha values were higher when estimated with items aggregated to class level.

This table shows the inter-item reliability statistics with two types of scale specifications: 1) when all available teaching methods are combined into a scale, and 2) when only student-centered methods are combined in a scale. It is evident that the

reliability statistics are quite close between the two types of specifications. However, one should note that when all items are being included in a scale, it measures something slightly different than when only student-centered methods are combined. One scale would represent instruction in general, whereas the other represents instruction using a specific approach, or a slightly different learning environment. The statistics presented in this table were calculated at the *individual student level*, where variability is far greater than at aggregated levels. It was important to test inter-item correlation at the level where it the real variability across measures is not deflated by regression to the mean, as would be the case if reliability analysis was applied to class aggregates on these teaching activities.

The conventional cutoff point for scale construction is a Cronbach's alpha of 0.5; all countries easily pass this level in science teaching methods, but not in mathematics. Because ultimately, the "snapshot analysis" is aimed at placing Kazakhstan in a context, in terms of its parameters on student-centered instruction, I only create a factor variable for the variables pertaining to science instruction, where the statistics are almost equally strong for all countries.

Table 6.7. Basic descriptives for the latent factor "Student-centered instruction in science"

Country	Eigenvalue of first factor	Proportion of total variance	Min value of factor	Max value of factor
Armenia	2.70	45%	-2.06	3.79
Georgia	3.22	54%	-2.31	4.74
Kazakhstan	3.81	63%	-2.58	2.64
Latvia	2.91	48%	-2.74	2.19
Lithuania	3.26	54%	-2.71	2.76
Russia	3.05	51%	-3.13	3.27
Ukraine	3.06	51%	-3.62	3.00

As Table 6.7 shows, across all countries, at least one latent factor accounting for roughly half of the variance of six items in the scale was identified. I use this factor as a single composite measure of student-centered instruction in regressions with science achievement outcomes as dependent variables. The principal components factor procedure automatically standardizes factor variables to a mean of 0 and standard deviation of 1, so that one unit on the factor scale corresponds to one standard deviation on in the level of exposure to student-centered instruction. The factor runs from roughly two to three standard deviations below the mean, to roughly two to three above the mean, although in Georgia the distribution appears to be highly skewed, with a long tail of high factor values..

The new variable – the composite factor for student-centered instruction in science – enters with two variables that fall in the category of "rote learning": memorization and independent work on science problems. These variables had acted as control covariates for the models shown in Table 6.4 and Table 6.5. The rest of the control covariates remain the same: student background, wealth and emotional comfort at school, school size, type of community, and at the third stage, the key teacher characteristics. The science composite factor has a main term and a square term, which, if significant, would be evidence of potentially nonlinear relationship.

Table 6.8 shows the regression coefficients on the composite factor, on its own as a sole predictor, and adjusted for covariates as described above. The point estimate on this factor for Kazakhstan is small (12 points for the main term), but statistically significant. The very small but still significant coefficient on the square term shows the presence of a U-shaped relationship, with a decline throughout most of the scale, and a slight increase at the end. At the same time, the rote learning methods contributed about 1/3 of a standard deviation in scores, controlling for background variables, and controlling for the presence of

student-centered instruction. In other countries, the composite factor showed no statistically significant coefficient, neither as a linear term, not with a square term. The other coefficients remained the same as in models whether the teaching activities entered as separate variables: strong and positive in Lithuania on both memorization and independent problem-solving, and strong and negative in Russia on memorization.

6.2.4. CONCLUSION: NO APPARENT ASSOCIATION BETWEEN STUDENT-CENTERED INSTRUCTION AND ACHIEVEMENT IN POST-SOVIET EURASIA

Based on these results, I conclude that *student-centered instruction in math and science is generally not associated with higher average scores* for students in the seven post-Soviet countries. Across countries, coefficients vary greatly, both in magnitude and in the direction of the sign, and no clear patterns could be discerned. Overall, the "rote learning" methods included as control variables in the models were more likely to show a large and statistically significant association, and to be less affected by the number and type of confounding covariates (i.e. student, school, and teacher background) included in the models. However, the magnitudes of the coefficients on such practices and, in rare cases, the direction of the relationships also varied. In Kazakhstan, group work on science project appeared to be related to higher scores, with a fairly large and robust coefficient. However, other experiment-related practices did not show statistically significant associations with achievement in science, which leaves us to wonder whether it is the group work per se that is more prevalent among high-achieving students, rather than the content of the projects around which it is organized. The composite scale variable, based on factor analysis of items measuring student-centered instruction in science showed *negative association with achievement* in science in Kazakhstan's data.

The cross-country analysis, however, has left out substantial depth in understanding the distribution of student-centered instruction across the different categories of students, as well as the school and teacher characteristics that may be associated with the types of instructional environments that the students are immersed in. This lack of depth was intentional: the purpose of this section was to look at how the teaching practices in Kazakhstan differ from the rest of the post-Soviet countries in the TIMSS group in terms of their association with achievement. The next section will examine Kazakhstan's data in greater depth, and apply more rigorous modeling and sampling specifications to mitigate selection bias and approach the estimation of the causal relationship between student-centered instruction and achievement.

Table 6.8. Regression coefficients using the composite factor variable "student-centered instruction in science".

Level of covariate adjustment	Classroom Activity	Armenia		Georgia		Kazakhstan		Latvia		Lithuania		Russia		Ukraine	
		coeff	se	coeff	se	coeff	se	coeff	se	coeff	se	coeff	se	coeff	se
Unadjusted	**Student centered instruction (composite factor)**	1.58	(7.15)	*-16.41*	*(4.37)*	*-12.20*	*(3.96)*	*-7.45*	*(1.84)*	*-9.04*	*(1.66)*	-0.14	(4.27)	-4.37	(2.84)
	||square term	0.40	(3.69)	*6.04*	*(1.18)*	3.87	(2.79)	*-3.66*	*(1.32)*	-1.28	(1.11)	3.56	(1.88)	-1.75	(1.70)
	Memorize science facts	19.90	(16.94)	16.64	(15.71)	*32.76*	*(12.63)*	*14.69*	*(7.00)*	*36.25*	*(7.50)*	*-29.37*	*(9.45)*	11.21	(11.75)
	Work science problems on own	1.35	(13.98)	11.10	(13.84)	*36.24*	*(16.99)*	*25.32*	*(12.33)*	*22.50*	*(7.48)*	*32.81*	*(9.61)*	21.40	(18.63)
Adjusted for student background, community size, and student social comfort at school	**Student centered instruction (composite factor)**	3.25	(6.35)	-9.41	(4.69)	*-13.94*	(3.35)	-2.30	(2.10)	*-4.16*	*(1.64)*	0.73	(4.28)	-1.21	(2.23)
	||square term	-1.66	(4.17)	2.21	(2.86)	4.28	(2.64)	-2.55	(1.33)	0.28	(0.96)	4.31	(1.85)	0.43	(1.33)
	Memorize science facts	35.64	(19.20)	5.43	(15.32)	*29.70*	*(11.43)*	9.03	(6.72)	*21.01*	*(6.41)*	*-27.78*	*(9.05)*	7.31	(9.28)
	Work science problems on own	1.48	(15.98)	2.55	(12.73)	*36.81*	*(13.60)*	8.99	(9.71)	*18.95*	*(6.76)*	*17.77*	*(8.77)*	1.87	(14.90)
Adjusted for student background, community size, student social comfort at school, and teacher characteristics	**Student centered instruction (composite factor)**	6.67	(6.64)	-9.47	(5.26)	*-12.48*	*(3.64)*	-3.67	(2.31)	-4.45	(2.13)	0.32	(3.92)	-1.75	(2.30)
	||square term	-2.58	(4.23)	5.13	(2.68)	*5.48*	*(2.79)*	-2.34	(1.78)	0.26	(0.97)	3.46	(2.01)	0.18	(1.43)
	Memorize science facts	29.39	(20.06)	13.59	(16.56)	*33.37*	*(10.93)*	9.16	(7.17)	*22.54*	*(6.89)*	*-26.98*	*(9.10)*	7.72	(8.79)
	Work science problems on own	-5.44	(15.31)	0.87	(15.52)	*32.45*	*(13.98)*	13.30	(9.94)	*26.38*	*(7.65)*	*24.38*	*(8.06)*	-2.03	(14.28)

Dependent variable: TIMSS science score (mean 500; sd 100).

6.3. STUDENT-CENTERED INSTRUCTION IN KAZAKHSTAN

In this section, I focus on examining the relationships between the use of student-centered methods of instruction in math and science in Kazakhstan. I begin with analyzing what factors are likely to predict the students' being placed in such instructional environments, and then proceed to restrict the sample through propensity score matching and run regression models with student-centered instruction as the treatment variable and TIMSS achievement scores as outcomes.

6.3.1. PREDICTING SELECTION FOR STUDENT-CENTERED INSTRUCTION

As I have noted above, one of the key considerations in analyzing observational (i.e. non-experimental) data, such as the student-reported level of exposure to a particular set of practices post-factum, is the potential for considerable selection bias in the sample. If the intensity of experience with a given practice or method is strongly associated with the background of the student or their school, it may be impossible to separate that influence from the association with instruction itself. For this reason, I begin my in-depth analysis of Kazakhstan's data by examining the distribution of student-centered teaching practices across the sample, comparing groups on their level of experience with such methods, based on their characteristics. Here, I follow my research questions:

1.b. What types of students and teachers are more likely to have been studying and teaching in student-centered environments?

2. What are the systemic and contextual factors associated with the use of student-centered methods in classroom instruction in math and science?

This section offers a quick look at the distribution of teaching practices depending on the proxies of socio-economic of the students. Table 6.9 demonstrates predictive models

of some of the student-centered methods in mathematics and science: for mathematics, being asked to perform measurement tasks in class, organizing information in tables, charts, or graphs, and working in small groups; in science, making observation and recording findings, conducting an experiment or investigation in science, and working in small group on a science experiment or investigation. The models are simple OLS regressions, adjusted for sampling weights and with clustering-robust standard errors.

Based on Table 6.9, I conclude that *while the distribution of student-centered instruction is not purely random, only a small handful of variables among those measured showed a statistically significant relationship with my instructional methods of interest*. Further, the pattern of the coefficients is also not stable, and that some variables contribute more than others to the likelihood of greater exposure to some methods of student-centered teaching, but not to others. There are no variables that invariably predict student-centered teaching across all methods. Among those that do predict greater likelihood of some methods are student immigration status (either student's own, or of parents); number of books in the house of the student; the experience of their teacher, resource level of the school (being low resourced); presence of a school policy for ability grouping for math and science lessons; a high proportion of economically disadvantaged students at the school (over 50%); Russian language of instruction, and the student's experience with bullying (inversely associated with student-centered instruction). Interestingly, teachers who had education as a second tertiary degree were less likely to engage in group work than teachers in with post-secondary non-tertiary education. The size of the community seems to not play a role in the type of instructional environment at the school, although smaller community size showed a large coefficient predicting the use of graphic organization in math lessons. Teacher age showed a large point estimate for the math-specific practices, but the standard errors were not sufficiently small to allow me to reject the null hypothesis.

In this *sample*, however, each three additional years of teacher age contributed nearly one full step on the scale of the math activities, which ran from 1 ("never") to 4 ("every lesson").

Consequently, there is *no evidence of significant bias in the choice of instructional practices* – or in the frequency with which they are applied in math and science lessons in Kazakhstan. A few variables seem to be showing a relationship, but it is far from a clear bias based on socioeconomic characteristics, or the type and location of community, or even the background of the teachers. There seems to be no clear structure or institutional path along which these practices would be implemented; it is more likely random than self-selected. However, an important caveat here is that as I pointed out above, these variables do not measure a single, well-defined policy or program, rather this is a post-hoc, student-reported and hence noisy measure of "reality" of instructional environments in fourth grade classrooms. Furthermore, even if all variables showed a lack of relationship with the methods of instruction that are being reviewed here, the absence of selection bias can never be empirically verified, and therefore, would remain an assumption.

As a next step, I change the approach to the measure of student-centered instruction in Kazakhstan, and examine it as a binary variable with two states: one, insufficient exposure to student-centered methods and predominantly rote learning methods; and two, regular exposure to student-centered methods, where rote learning methods may or may not be equally present. I then construct two "treatment groups" corresponding to the two states, and further restrict the sample to have the groups match on the estimated propensity score (see Chapter 4 for methodology on propensity scores).

Table 6.9. Predicting exposure to selected student-centered teaching methods based on background characteristics.

		MATH METHODS						SCIENCE METHODS					
	Variable	Measurement		Tables charts graphs		Group work		Sci Observation		Sci doing experiments		Sci group work on	
		coeff	se	coeff	se	coeff	se	coeff	se	coeff	se	coeff	se
Student	Having a calculator	-0.027	(0.02)	0.046	(0.05)	0.054	(0.05)	-0.049	(0.04)	-0.066	(0.03)	0.020	(0.05)
	Having a computer	0.036	(0.02)	0.002	(0.04)	-0.006	(0.04)	0.035	(0.04)	0.078	(0.04)	0.084	(0.05)
	Number of books in the house	0.039	(0.01)	0.048	(0.02)	0.019	(0.02)	0.029	(0.02)	0.030	(0.02)	0.071	(0.02)
	Student born in country	-0.026	(0.05)	0.072	(0.06)	0.031	(0.06)	0.026	(0.06)	0.067	(0.07)	0.084	(0.09)
	Both parents born in country	0.059	(0.06)	0.029	(0.06)	0.129	(0.05)	0.059	(0.06)	0.049	(0.06)	0.033	(0.07)
	Russian language	0.155	(0.10)	0.119	(0.10)	0.300	(0.08)	-0.035	(0.10)	-0.231	(0.10)	0.072	(0.10)
	Was made fun of at school	-0.083	(0.03)	-0.101	(0.05)	-0.035	(0.05)	-0.098	(0.05)	-0.115	(0.04)	-0.056	(0.05)
Teacher	Age of teacher	0.290	(0.17)	0.226	(0.22)	0.437	(0.23)	0.090	(0.29)	0.130	(0.27)	0.167	(0.27)
	Experience of teacher	0.009	(0.02)	-0.009	(0.02)	-0.032	(0.02)	0.015	(0.02)	0.011	(0.02)	-0.007	(0.02)
	Male teacher	0.013	(0.09)	-0.234	(0.15)	-0.235	(0.10)	0.084	(0.12)	-0.022	(0.11)	-0.125	(0.14)
	Education ISCED 5A First degree	0.058	(0.07)	-0.023	(0.10)	-0.004	(0.11)	0.011	(0.11)	0.018	(0.10)	-0.003	(0.12)
	Education ISCED 5A Second degree	-0.081	(0.07)	-0.120	(0.12)	-0.252	(0.12)	-0.060	(0.11)	0.000	(0.10)	-0.103	(0.12)
	Teacher experience square term	0.000	(0.00)	0.000	(0.00)	0.001	(0.00)	0.000	(0.00)	0.000	(0.00)	0.000	(0.00)
	Teacher age square term	-0.040	(0.02)	0.004	(0.04)	-0.047	(0.04)	-0.021	(0.04)	-0.031	(0.04)	-0.026	(0.04)
	Collaborate with others - discuss concepts	0.110	(0.04)	0.006	(0.05)	0.001	(0.05)	0.053	(0.06)	0.042	(0.05)	0.028	(0.06)
	Collaborate with others - prepare for lessons	-0.038	(0.03)	0.110	(0.05)	0.067	(0.05)	0.024	(0.05)	-0.035	(0.05)	0.003	(0.05)
School	School total enrollment	0.000	(0.00)	0.000	(0.00)	0.000	(0.00)	0.000	(0.00)	0.000	(0.00)	0.000	(0.00)
(ref: community >500K)	Community 100001 TO 500000 PEOPLE	-0.022	(0.12)	0.015	(0.13)	-0.019	(0.11)	0.160	(0.13)	0.035	(0.11)	0.048	(0.15)
	Community 50001 TO 100000 PEOPLE	0.022	(0.13)	-0.190	(0.20)	0.090	(0.15)	0.323	(0.19)	-0.082	(0.16)	0.109	(0.21)
	Community 15001 TO 50000 PEOPLE	-0.096	(0.13)	-0.328	(0.16)	-0.058	(0.15)	0.102	(0.16)	0.073	(0.16)	0.084	(0.16)
	Community 3001 TO 15000 PEOPLE	-0.022	(0.12)	-0.072	(0.15)	0.082	(0.16)	0.227	(0.15)	0.109	(0.13)	0.072	(0.16)
	Community 3000 PEOPLE OR FEWER	-0.026	(0.16)	0.047	(0.17)	0.126	(0.17)	0.148	(0.18)	0.077	(0.17)	0.121	(0.21)
ref: <10%	Percent poor students 11-25%	-0.081	(0.06)	-0.017	(0.12)	-0.135	(0.11)	-0.183	(0.10)	-0.046	(0.09)	-0.139	(0.11)
	Percent poor students 26-50%	0.078	(0.13)	-0.090	(0.10)	-0.149	(0.11)	0.010	(0.12)	-0.020	(0.12)	-0.054	(0.13)
	Percent poor students more than 50%	-0.274	(0.16)	-0.104	(0.29)	-0.135	(0.18)	-0.483	(0.25)	-0.758	(0.28)	-0.624	(0.26)
ref: Yes	School ability grouping: NO	-0.101	(0.07)	-0.155	(0.09)	-0.275	(0.10)	-0.084	(0.09)	-0.064	(0.08)	-0.230	(0.10)
Ref: High	Index of resource availability at school: Medium	0.106	(0.06)	-0.002	(0.08)	0.124	(0.09)	*-0.257*	*(0.10)*	-0.189	(0.09)	-0.114	(0.10)
	Index of resource availability at school: Low	-0.109	(0.16)	-0.363	(0.13)	-0.406	(0.15)	-0.431	(0.18)	-0.519	(0.15)	-0.228	(0.21)
	Constant	*1.194*	*(0.47)*	*1.845*	*(0.47)*	*1.856*	*(0.42)*	*2.464*	*(0.60)*	*2.442*	*(0.53)*	*2.318*	*(0.51)*

6.3.2. STUDENT-CENTERED INSTRUCTION IN KAZAKHSTAN AS A BINARY VARIABLE

The hypothesis behind the construction of the binary variable is the following: the true effect of student-centered instruction on achievement outcomes is positive, but it is not linear or even continuous; rather, it is necessary to establish a routine where methods would be combined and used concurrently, so that the entire environment changes, rather than a given practice. You are either "in" or "out": either you do practice or engage in student-centered instruction, or you do not. To test this hypothesis, I employ two methods. In one method, I "manually" create indicator variables for student-centered instruction in math and science, based on an *intuitive* coding of the items that would normally enter such a combined set of practices. In other words, I code the indicator variables in a logical, intuitive way, albeit with reference to the literature on student-centered instruction (Bransford et al., 2004), since there is no single "recipe" that would guarantee a "sufficient" level of engagement. This specification is explained in Chapter 5, Table 5.4 for the student-centered instruction in math and Table 5.5 for student-centered instruction in science.

In the other approach, I use the slightly more automated way of coding student-centered instruction into a binary treatment variable, using factor analysis. I test an alternative specification for science instruction, based on the continuous factor variable estimated using the six items representing student-centered instructional methods, as shown in Table 5.2. This factor, roughly normally distributed across the sample is simply divided in half: those who received such instruction at a frequency level below the mean were coded as "controls", and those above the mean were coded as "treated". I perform this additional specification to test the potential bias I may have introduced with what is arguably an arbitrary specification of who should be considered sufficiently "exposed", or

"treated". Breaking a composite factor at the mean may be seen as a more "objective" way of coding the treated vs. controls division, albeit still resting on an assumption that the mean level of exposure is the magic level of "sufficiency". This additional approach to coding science instruction produced a treatment group consisting of 48% of the sample. This approach to coding the treatment variable in science based on the composite factor is referred to as Coding II, as opposed to Coding I, which is based on my judgment of what combination of methods is closest to the literature on student-centered instruction. The approaches to coding the treatment variables are presented in Chapter 5. As noted above, due to the weak inter-reliability of the four math instruction items, no factor-based specification was done for math.

6.3.3. MATCHING ON THE PROPENSITY SCORES

Once the Kazakhstan variables were coded into treatment dummies, I then proceeded to estimate the propensity score models for the groups divided by these binary variables. As explained in Chapter 5, the propensity score models are probit regressions estimated with clustering-robust standard errors. Propensity scores are the predicted probabilities of treatment, based on these models. As could be expected given the earlier analysis of predictive models of continuous student-centered instruction through a series of covariates (Table 6.9), very few of the variables entered in propensity score models were statistically significant. However, formal hypothesis testing is not crucial when estimating an algorithm for matching on the propensity score. What is more important is reaching an adequate estimation of the probability of treatment, so that, once that probability is adjusted for, one can be certain that nothing else can possibly affect the estimate of the treatment association with the outcome. In reality, however, this is an untestable proposition, since in post-hoc observational studies the possibility of an unmeasured factor

interfering in the treatment assignment can never completely be rejected. Therefore, achieving acceptable balance on key covariates in a matched sample serves as an indication, but never as complete proof, that treatment assignment is "ignorable" (Rosenbaum & Rubin, 1983).

The propensity score models estimate the probability of receiving "treatment", or sufficient exposure to student-centered pedagogy, based on background characteristics of students, teachers, and schools, as well as a multitude of interactions between these variables, their higher-order terms, and logarithms. The propensity score models were repeatedly estimated, with close to one hundred models fit in order to reach the best balance between the groups on the student, teacher, and school variables. During this iterative process, the groups are examined after each propensity score model is run and units (i.e. students) are matched on the probabilities of treatment based on that model, and one seeks to determine that there exists no significant difference between the two treatment groups.

The matching on the propensity scores is accomplished through the PSMATCH2 algorithm in Stata. Each treatment unit is matched with nearest neighbor on the propensity score, and control units with no matches in the treated group are excluded from the matched sample. Units are then assigned frequency weights corresponding to the number of times they should be used in the analysis in order to have sufficient balance in the matched sample: treated units are assigned a weight of 1 and control units are assigned weights of 0 if not matched, or equal to the number of times they were matched. As I noted above, matching was performed numerous times with various specifications, and each time I examined balance upon matching to find the best possible algorithm.

Table 6.10. Variables in the final propensity score model for student centered instruction in science

Variable name	Label	Type
as4gsex	gender of student	Binary
studage	age of student	Continuous
russian	language of instruction	Binary
stspeaklang	speaking language of instruction at home	Binary
nativeborn	native to Kazakhstan	Binary
perc_povschool	percent poor students at school	Proportion
rel_wealth	interaction between SES of student and percent poor students at school	Continuous
hv_calc	having a calculator	Binary
wealthy	wealthy possessing more than 4 items of home possessions	Binary
books_home	number of books at home	Ordinal, 4-point
books_wealth	interaction between books and wealthy	
tchprep_edsci	teacher trained in education and science	Binary
tchprep_edmath	teacher trained in education and math	Binary
teachexp	teacher experience	Continuous
squteachexp	teacher experience square term	Continuous
community	community size	Ordinal, 6 point
tchtrn2	teacher trained ISCED 5a 4yr college	Binary
tchtrn3	teacher trained ISCED 5a 5-6yr univ or higher	Binary
clsize_cat	class size as category	Ordinal, 4 point
clsize_raw	class size number	Continuous
st_sciwkown	science working on sci problems on own	Ordinal, 4 point
st_scimemorize	science memorizing how to work problems	Ordinal, 4 point
sci_hwk	amount of homework in science	Ordinal, 4 point
tradteach_hmwk	interaction traditional methods and homework	Interaction
ac4sgasc	school does ability grouping in science	Binary
lnschoolsize	log school size	Log:continuous nonnegative

As a result of these multiple iterations, two models were found to be producing the best possible balance on important confounding covariates in math and science. These models are presented in Table 6.10 and Table 6.11. *It is important to note that a much larger array of variables, their higher order terms, and interactions, were tested for propensity score models.* In the process of checking balance produced by such models, variables were added and removed from the model, until a reasonably good balance was found,

understanding the relative importance (or unimportance) of some covariates on the final outcome.

Table 6.11. Variables in the final propensity score model for student-centered instruction in math

Variable name	Label	Type
studage	student age	Continuous
russian	russian speaking	Binary
nativeborn	native to Kazakhstan	Binary
tchprep_ed~i	teacher trained in education and science	Binary
tchprep_ed~h	teacher trained in education and math	Binary
tchprep_ma~i	teacher trained in math and science	Binary
_Iwealthy_1	wealthy index (saturated, each category entered as a dummy)	Set of binaries
IrusXwea~1	interactions between language of instruction and wealthy status	
compuse_both	computer use both in school and at home	Binary
hv_internet	internet connection at home	Binary
books_home	books in the home categories (1="none", 5="three or more bkcases")	Ordinal, 5-point scale
books_wealth	interaction books and wealthy	
perc_povsc~l	percent poor students at school	Proportion
rel_wealth	interaction student's own wealth and percent poor at school	
bullied	index of being bullied at school	Ordinal, 1-5 scale
teachexp_yrs	teacher experience (years)	Continuous
squteachexp	teacher experience square term	Continuous
_Iatdgcoll_2	interactions with colleagues index - medium	Binary
_Iatdgcoll_3	interactions with colleagues index - high	Binary
tchtrn2	teacher trained ISCED 5a 4yr college	Binary
tchtrn3	teacher trained ISCED 5a 5-6yr univ or higher	Binary
teachcond	index teacher perception of work conditions (infrastructure, repairs, supplies; 1 “Not a problem”, 2 “Minor problem”, 3 “Major problem”)	Ordinal, 3-point scale
_Iac4gchts_2	principal perception of teacher satisfaction - medium	Binary
_Iac4gchts_3	principal perception of teacher satisfaction - high	Binary
clsize_cat	class size as category	Ordinal, 4-point scale
clsize_raw	class size as number of children	Continuous
_Icommunit~2	community size (saturated, each category is a dummy)	Set of binaries
lnschoolsize	log school size	Log, nonnegative continuous
ac4mgamc	school groups by ability for math	Binary

Balance diagnostics for the two models selected for propensity score matching are presented in Table 6.12 and Table 6.13: the means and standard deviations on variables in each model are shown before and after matching. Balance improved on most parameters

after matching, both on the means and standard deviations, which indicates that the groups became closer on background characteristics than if taken without propensity score matching. Even when most differences in means and standard deviations prior to matching were small, the potential for eliminating unobserved bias is greater when the distributions of observed variables are closer between the two groups, and hence the estimate of the association between the treatment and the outcome is more robust in a matched sample.

Table 6.12. Balance on key covariates before and after matching, MATH.

Variable	Sample	Mean		Standard Deviation	
		Treated	Control	Treated	Control
studage	Unmatched	10.647	10.573	0.5	0.5
	Matched	10.647	10.689	0.5	0.6
as4gsex	Unmatched	1.483	1.5133	0.5	0.5
	Matched	1.483	1.4905	0.5	0.5
russian	Unmatched	0.46591	0.3996	0.5	0.5
	Matched	0.46591	0.3952	0.5	0.5
nativeborn	Unmatched	0.91288	0.92354	0.3	0.3
	Matched	0.91288	0.94003	0.3	0.2
tchprep_ed~i	Unmatched	0.88699	0.84043	0.3	0.4
	Matched	0.88699	0.91225	0.3	0.3
tchprep_ed~h	Unmatched	0.04735	0.08511	0.2	0.3
	Matched	0.04735	0.03725	0.2	0.2
tchprep_ma~i	Unmatched	0.00758	0.04455	0.1	0.2
	Matched	0.00758	0.01705	0.1	0.1
tchprep_ot~r	Unmatched	0.00947	0.00997	0.1	0.1
	Matched	0.00947	0.01199	0.1	0.1
wealthy	Unmatched	4.7727	4.3903	2.2	2.2
	Matched	4.7727	4.8188	2.2	2.2
compuse_both	Unmatched	0.24053	0.13564	0.4	0.3
	Matched	0.24053	0.30619	0.4	0.5
hv_internet	Unmatched	0.27841	0.23404	0.4	0.4
	Matched	0.27841	0.30177	0.4	0.5
books_home	Unmatched	2.8081	2.4501	1.2	1.2
	Matched	2.8081	2.7014	1.2	1.2
books_wealth	Unmatched	14.263	11.686	10.1	9.6
	Matched	14.263	13.856	10.1	9.9
perc_povsc~l	Unmatched	1.6875	1.6496	0.9	0.8
	Matched	1.6875	1.8668	0.9	0.9
bullied	Unmatched	0.03909	-0.04214	0.8	0.8

Table 6.12., continued

	Matched	0.03909	0.04602	0.8	0.7
teachexp_yrs	Unmatched	18.809	17.461	9.1	9
	Matched	18.809	19.138	9.1	8.6
ac4gchts	Unmatched	2.2961	2.2533	0.5	0.6
	Matched	2.2961	2.2431	0.5	0.6
tchtrn2	Unmatched	0.43624	0.40891	0.5	0.5
	Matched	0.43624	0.4798	0.5	0.5
tchtrn3	Unmatched	0.30492	0.26263	0.5	0.4
	Matched	0.30492	0.25	0.5	0.4
teachcond	Unmatched	-0.25436	0.07265	0.8	0.8
	Matched	-0.25436	-0.32017	0.8	0.7
clsize_cat	Unmatched	1.952	1.9328	0.5	0.5
	Matched	1.952	1.827	0.5	0.5
clsize_raw	Unmatched	25.129	24.219	5.3	5.5
	Matched	25.129	23.924	5.3	5.1
community	Unmatched	3.9129	3.8471	1.7	1.7
	Matched	3.9129	3.9539	1.7	1.6
lnschoolsize	Unmatched	6.5697	6.691	0.7	0.8
	Matched	6.5697	6.6462	0.7	0.7
ac4mgamc	Unmatched	1.4792	1.6144	0.5	0.5
	Matched	1.4792	1.5158	0.5	0.5

Table 6.13. Balance on key covariates before and after matching, SCIENCE

		Mean		**SD**	
Variable	Sample	Treated	Control	Treated	Control
studage	Unmatched	10.671	10.546	0.5	0.5
	Matched	10.671	10.656	0.5	0.5
stspeaklang	Unmatched	0.80093	0.83896	0.4	0.4
	Matched	0.80093	0.7743	0.4	0.4
nativeborn	Unmatched	0.93076	0.91003	0.3	0.3
	Matched	0.93076	0.92876	0.3	0.3
perc_povsc~l	Unmatched	1.6325	1.8061	0.9	0.9
	Matched	1.6325	1.5752	0.9	0.8
rel_wealth	Unmatched	7.3848	6.857	5.4	4.9
	Matched	7.3848	7.3822	5.4	4.9
hv_calc	Unmatched	0.9534	0.91003	0.2	0.3
	Matched	0.9534	0.96804	0.2	0.2
wealthy	Unmatched	4.6178	4.0243	2.2	2.2
	Matched	4.6178	4.8123	2.2	2.2
books_home	Unmatched	2.7856	2.3446	1.2	1.1
	Matched	2.7856	2.773	1.2	1.2

Table 6.13., continued

books_wealth	Unmatched	13.702	10.425	9.9	9.2
	Matched	13.702	14.182	9.9	10.3
tchprep_ed~i	Unmatched	0.91744	0.87854	0.3	0.3
	Matched	0.91744	0.92277	0.3	0.3
tchprep_ed~h	Unmatched	0.06658	0.07197	0.2	0.3
	Matched	0.06658	0.06525	0.2	0.2
teachexp	Unmatched	3.8928	3.5394	1.4	1.5
	Matched	3.8928	3.8888	1.4	1.4
teachexp_yrs	Unmatched	19.255	16.755	8.9	9.1
	Matched	19.255	19.18	8.9	8.8
squteachexp	Unmatched	450.56	364.34	369.3	339.7
	Matched	450.56	445.58	369.3	354.9
community	Unmatched	3.9454	3.5497	1.6	1.8
	Matched	3.9454	4.008	1.6	1.8
russian	Unmatched	0.51598	0.31219	0.5	0.5
	Matched	0.51598	0.52929	0.5	0.5
tchtrn2	Unmatched	0.40613	0.41925	0.5	0.5
	Matched	0.40613	0.41811	0.5	0.5
tchtrn3	Unmatched	0.26698	0.2879	0.4	0.5
	Matched	0.26698	0.27097	0.4	0.4
clsize_cat	Unmatched	1.9048	1.9478	0.5	0.5
	Matched	1.9048	1.9121	0.5	0.5
clsize_raw	Unmatched	24.381	24.705	5	5.5
	Matched	24.381	24.183	5	5.4
st_sciwkown	Unmatched	3.7383	3.5947	0.6	0.7
	Matched	3.7383	3.7144	0.6	0.6
st_scimemo~e	Unmatched	3.5732	3.2785	0.7	0.9
	Matched	3.5732	3.5526	0.7	0.7
sci_hwk	Unmatched	2.5486	2.5376	0.8	0.8
	Matched	2.5486	2.5393	0.8	0.8
tradteach_~k	Unmatched	34.439	30.343	15.4	15.2
	Matched	34.439	33.88	15.4	14.7
ac4sgasc	Unmatched	1.6605	1.6419	0.5	0.5
	Matched	1.6605	1.6198	0.5	0.5
lnschoolsize	Unmatched	6.6082	6.5678	0.6	0.8
	Matched	6.6082	6.6513	0.6	0.8
as4gsex	Unmatched	1.474	1.5065	0.5	0.5
	Matched	1.474	1.5047	0.5	0.5

Once the specifications of probability models for student-centered instruction in math and science were finalized, and propensity scores estimated, I turn to the estimation of the association between the treatment and the outcomes. There are two approaches for using the propensity scores for bias control, both of which enhance the plausibility of causal inference. One method is *regression-adjusted matched estimates*, and the other *is weighting on the propensity scores*.

6.3.4. ESTIMATING THE TREATMENT EFFECT: REGRESSION - ADJUSTED MATCHED ESTIMATES

For a regression-adjusted matched estimate, I run a regression of the outcome on the treatment variable, using frequency weights calculated by the PSMATCH2 algorithm in STATA. This algorithm performs nearest-neighbor matching of the treated and control observations, and assigns frequency weights to control observations equal to the time they were determined "nearest neighbors" to a treated observation. Treatment observations are always assigned a weight of 1. This method, called "matching with replacement" allows for a control observation to be used several times in an estimation procedure, if it happens to be a good match for more than one treated unit. However, this algorithm discards control observations for which no treated counterfactuals could be found, thereby reducing the sample size and the representativeness of the sample relative to the overall population.

Finally, I perform additional control of the covariate effects, by entering into the outcome regression all variables used to estimate the probability of receiving student-centered instruction. This ensures that if some imbalance remained in background characteristics after matching, it is removed from the estimate of the relationship between student-centered instruction and achievement outcomes. Thus, the final outcome models are OLS regressions of the five plausible values in math or science on the "treatment" (the

binary variable for student-centered instruction), weighted using the frequency weights estimated by the propensity score models, as follows:

SCORE (Math, Sci) = Cons. + Treat(Math, Sci) + Covariates [PSWEIGHT]

- where *SCORE* is the five TIMSS plausible values in math or science, regressed on the dummy treatment variable representing student-centered instruction in math or science, respectively, plus a set of covariates used in the estimation of propensity scores, adjusted for matching using propensity scores PSWEIGHT. As in models estimating effects in non-matched samples, the jackknifing procedures adjusting the estimation for clustered sampling and for the imputation error, specified by IEA for use with TIMSS test plausible values, were applied in propensity score matched regressions. This ensures that standard errors are robust, and minimizes the possibility of incorrect rejection of null hypothesis (Type I error). The full output of the regressions is presented in Appendix A6.

6.3.5. ESTIMATING THE TREATMENT EFFECT: WEIGHTING ON THE PROPENSITY SCORE

The *propensity score weighting* method for estimating treatment effects is not as restrictive as *matching* on the propensity score, where control observations are dropped from the sample if there are no treated observations that can be matched with them, but it allows for a stricter control of selection bias than a simple regression on an unmatched sample. In this method, for the estimation of average treatment effect treated observations are assigned weights equal to the inverse of their estimated propensity score (or, in other words, their probability of receiving treatment), and control observations are assigned weights equal to the inverse to their propensity score subtracted from 1 (or, simply put, the probability of their not receiving treatment), as follows:

Weight (Treated) = 1/e(x)

Weight (Control) = 1/(1-e(x))

- where **e(x)** is the estimated propensity score (Gelman & Hill, 2007).

Because no observations are dropped from the sample, statistical power remains as strong as it was prior to matching, while the construction of the weights ensures that the influence of extreme observations, which are not likely to have matched counterfactuals, is minimal. For estimating the average treatment effect (ATE), one simply regresses the outcome of interest on the treatment variables and appropriate control variables, and assigns the calculated propensity score weights as probability weights for the regression. The point estimates on the treatment and control variables can be expected to change, as the weighed sample still includes all observations, and each observation is generally used only once (as opposed to repeated use of matched controls in the regression-adjusted matched estimation shown above).

6.3.6. MULTILEVEL MODELING

As an additional means of accounting for the nested structure of the sample, I used HLM 6.08 multilevel modeling software (Bryk, Raudenbush, & Congdon, 2010) to fit the outcome models for math and science. The outcomes – the TIMSS learning scores – entered as Level 1 outcome variables, along with a set of student characteristics, such as age, gender, nativity to Kazakhstan, a set of SES proxies (number of books in the house, home possessions, wealthy index), and negative social experience in the school (i.e. bullying), all of which were Level 1 controls, with group-mean centering (Raudenbush & Bryk 2002). The treatment variables – the binary indicators for student-centered instruction in math and science, coded for propensity score matching – were situated at Level 2., uncentered.

Besides the treatment variables, Level 2 included relevant covariates for teacher characteristics, including their age, experience (with a square term), specialization in training, and level of formal education completed; and school level variables, including school enrollment and percent of economically disadvantaged students. Because in the Kazakhstan's TIMSS sample nearly all schools had only one teacher and one class, and only a few schools had two TIMSS classrooms, a three-level estimation (i.e. student, teacher, school) was not feasible, and therefore I combined both teacher and school variables into the Level 2 file.

Adjustment for propensity score matching was achieved in a similar way as described above: frequency weights calculated by the propensity score matching algorithms were read as frequency weights for the model. Because the propensity score models estimate the probability for each student to receive student-centered instruction, these frequency weights were placed at the student level in the HLM framework, even as the treatment variables themselves were placed at Level 2. In addition to the propensity score weights, I placed the TIMSS survey teacher weight as the Level 2 weight, as recommended by IEA (2008).

The ability to use weights at different levels and to correctly account for imputation uncertainty surrounding plausible values are the two benefits of HLM as a package for multilevel modeling that are not available in general-use software, such as SPSS's MIXED command, or Stata's XTMIXED. The plausible values for math and science were entered as outcomes in models with the math and science treatment variables, respectively. Therefore, the estimation of error variance properly controls for the clustering of students into schools, the imputation uncertainty, and the adjustment for the probability of receiving

student centered instruction, estimated using propensity score matching as described above.

6.3.7. RESULTS OF PROPENSITY SCORE ADJUSTED MODELS

The results of the propensity score matched, propensity score weighted, and propensity-score matched HLM models estimations of the effect of student-centered instruction on math and science achievement outcomes are presented in Table 6.14. As the table shows, the point estimates are negligible, and standard errors are large to the point that no statistical significance can be attained even at levels far more generous than the conventional 5%: errors across the matched and weighted estimates are close in magnitude to the estimates themselves, and the estimates are barely 0.1 of a standard deviation in the learning scores.

Table 6.14. Treatment coefficient estimates from propensity score adjusted models.

	Matched		Weighted		HLM	
	coeff	**se**	**coeff**	**se**	**coeff**	**se**
Math ("intuitive coding")	10.07	(9.42)	5.50	(8.50)	7.61	(13.74)
Science: Coding 1 ("intuitive coding")	10.22	(8.21)	7.41	(8.12)	8.05	(12.50)
Science: Coding 2 (based on continuous factor)	11.20	(8.22)	8.47	(7.67)		

Note: Cell entries are regression estimates, standard errors in parentheses.
Dependent variable: Five TIMSS plausible values, in math and science respectively.

The weighted point estimates are substantially lower than the matched estimates, indicating that the averaging out the estimate across all observations, including unmatched controls, serves to dampen the association. In causal inference language, which is used here illustratively, this means that had the students in the control group been exposed to student-centered instruction at the same level as the treated students, the differences in

scores that we would have observed would have been even lower than the average results of roughly 10 test score points.

The standard errors are clustering robust in all cases – in the single-level models, using the clustering adjustment variables provided by the TIMSS dataset, and in multilevel models, through explicit accounting for the nested structure in the model itself. The results of the multilevel models have especially large standard errors, and point estimates somewhat lower than those from a single-level matched regressions. These standards add another level of robustness to the error variance estimation, and therefore, are perhaps close to the true variance around these variables in the population.

In sum, the conclusion derived from the results in this table is the same as in the sections above: there is *no evidence of an association between student-centered instruction in math or science and student achievement*, neither positive nor negative. Point estimates on the composite binary coding of student-centered instruction are decidedly lower than those estimated using variables directly, without collapsing or recoding them, which indicates that even in this sample alone, the differences in math and science scores are not associated with a "sufficient" level of exposure, but is perhaps truly nonlinear, without a consistently positive or negative link with achievement outcomes.

The control of covariates proved to be an extra level of bias control, but one that decidedly changed the results of the estimations one way or another. Propensity score adjustment ensured that the students who received student-centered instruction were matched with those in the more traditional classrooms, so that the absence of an association cannot be attributed to selection bias – at least not on observed variables. However, where slight selection bias was present, it was the higher achieving students that were more likely to be exposed to student-centered methods. Therefore, bias control predictably did not

increase the magnitude of the learning score differences, but quite to the contrary, it reduced them to a negligible level, which even if it were statistically significant, would hardly be of substantive interest in terms of explaining score gains.

6.4. CONCLUSION

6.4.1. SUMMARY OF CHAPTER

In this chapter, I used quantitative methods to explore whether student-centered instruction in math and science improves student achievement outcomes in these cognitive areas, controlling for other relevant factors, and in comparison with traditional rote learning methods of instruction. The data for this analysis came from the TIMSS 2007 student achievement study, which was administered under the auspices of the International Association for the Evaluation of Educational Achievement (IEA) in 67 locations around the world, including national and subnational participants. While I was interested specifically in estimating the score differences associated with student-centered instruction in Kazakhstan, I began with an analysis of data across seven post-Soviet countries, which I hypothesized to be quite similar in their education systems to Kazakhstan, due to their common legacy. In these seven countries, I examined the distributions of relevant teaching practices, and fit several regression models with TIMSS learning scores as dependent variables, and the student-centered methods as predictors of interest, controlling for key covariates in student background, teacher characteristics, and school environment. This cross-country analysis showed substantial variation of coefficient sizes on teaching practices across the seven post-Soviet states. Not only the coefficients and standard errors varied, but also the direction of the estimated relationship. Composite variables reflecting a common latent variable across all of the student-centered practices similarly did not show a

statistically significant association or a consistent pattern across countries (a small negative association was found in Kazakhstan's dataset). In sum, I saw *no evidence that my hypothesis was true* and student-centered instruction had a positive association with achievement outcomes in math and science in post-Soviet countries.

As a next step, I examined the Kazakhstan TIMSS dataset, looking for indications of potential selection bias, through predictive models where a few of the designated student-centered methods of instruction were regressed on student, teacher, and school characteristics. Only a few variables showed statistically significant relationships with the frequency of exposure to student-centered instruction, and there was no clear pattern of bias, at least on observed measures.

Pursuing a hypothesis that perhaps, only sufficiently regular exposure to student-centered instruction could be expected to show an association with achievement outcomes, I created binary treatment variables for math and science student-centered instruction. Students were divided into two groups: the "treated", or those that received a sufficient, by my definition, level of such teaching, and the "control", designating those students whose exposure to student-centered instruction was minimal. In addition to my own definitions of "sufficiency" in treatment based on literature (shown in Chapter 5), I created an alternative binary treatment variable for science instruction, by simply cutting in half the composite factor created with all six science teaching practices (since the factor variable was standardized, the mean was set at zero). Despite weak indications of selection bias, I still proceeded to impose tighter controls on the Kazakhstan sample, through propensity score matching, where I fit multiple models estimating the probability of being in the treatment group as a function of student, teacher, and school characteristics. I finally selected the

models that, when used to match the samples on the estimated probability, produced the best possible balance on all key covariates.

Finally, I used the frequency weights estimated by the matching algorithms, as well as the probability weights calculated with estimated propensity scores, to estimate the treatment coefficients in math and science on student achievement outcomes. I used both the regression-adjusted matched modeling recommended by Gelman and Hill (2007), as well as weighting on the propensity scores, and fit both the single-level models (which, nonetheless, correctly account for the interdependence of errors due to nesting of students within classrooms and schools, and for the imputation uncertainty surrounding the outcome learning scores), and multilevel, or hierarchical linear models with two levels: 1) student level and 2) teacher & school level.

Once again, upon closer examination of Kazakhstan's data, *student-centered instruction showed no significant association with student achievement, neither in math nor in science*. Clustering-robust standard errors were almost as large as point estimates, and point estimates themselves were barely of substantive interest due to their negligible magnitude relative the scale of the outcome variables.

6.4.2. INTERPRETATION OF RESULTS

As these results showed, *my hypothesis of a positive effect of student-centered instruction did not prove to be correct*: results are inconclusive and error variance is too high to reliably estimate the relationship between the teaching practices of interest and achievement outcomes in math and science. Even without considering the statistical significance of sample findings, the point estimates are very small, particularly when

combined to form a proxy of an underlying approach to teaching as a whole, as compared with individual items.

There can be several potential explanations for this lack of an association. First, as I pointed out above, the data are non-observational: they were collected from student self-reports on the TIMSS student questionnaire, and therefore contain some degree of measurement error. Specifically, while appropriate efforts were made to minimize noise surrounding student perceptions of classroom teaching – by aggregating responses to a class mean – it is possible that factors not related to the actual practice of instruction in the classroom may have affected the responses of groups of students. In answering survey questions about the frequency of a given activity or practice in their lessons, students were asked to recall their experiences over the past academic year, which may had led students to different conclusions about the same set of practices. In addition, the measures themselves may not be adequate reflections of student-centered teaching as a complex phenomenon. The items that were available as proxies of student-centered instruction – group work, emphasis on experiment, requirements to reflect on the material covered in class, to organize written information, etc. – may not adequately capture the width of practices in a child-centered classroom, or may not draw the distinction sufficiently well between *student-centered* instruction as a philosophically different approach to teaching, and the quantity of instruction per se.

Another potentially important caveat and a possible explanation for inconclusive results is the lack of information on the *quality* of teaching in student-centered instructional methods. In other words, it is well known that teachers vary substantially in the ways that they use what may seem as a standard set of practices (such as, in traditional instructional environments, lecture, test preparation, or cold-calling), and yet such data is rarely

measured on large enough scale to be used in combination with student performance data[2]. In the analysis described in this chapter, many of the key qualifications of teachers are held constant, including their gender, age, experience, education level, and specialization in training, as well the extent of their collaboration with peer teachers. However, while these are important predictors of teacher quality, they are far from determinants of the teachers' instructional effectiveness specifically *in student-centered teaching practices*. In other words, teachers may possess the required diplomas but lack actual knowledge or skills. Alternatively, they may be well qualified in their subject areas, and still not be effective teachers in general.

Further still, teachers may be both competent and effective in teaching their subjects, but lack the desire, the confidence, or the skills to effectively use student-centered teaching methods, which only recently have begun to enter the mainstream debate on effective teaching in the post-Soviet region, despite their long history in frontline pedagogy. In a region still governed by centralized policy and with all-encompassing national public education systems, the directives and leadership from state education authorities is crucial. A vacuum of leadership in instructional reform can contribute to increasing variation in the quality of teaching, and to confusion as to how, if at all, the non-traditional teaching methods are to be taken and implemented by schools. The case study that I undertook as part of this dissertation, described in Chapter 7, sheds some light on this aspect in the context of Kazakhstan.

Finally, one should not reject the potential explanation that in fact, *student-centered instruction may well be ineffective* when it comes to attaining hard-core achievement

[2] In TIMSS 1995, such a large scale effort took place: teachers were observed and videotaped; video data later analyzed for effective teaching research.

outcomes, such as mathematics and science test scores. These practices may be beneficial for students in a variety of ways beyond these narrowly defined outcomes. They may (or may not) help their social and emotional development, their ability to engage with complex subject matter and build self-confidence in approaching academic and social challenges, learning to follow a plan while working in teams, or realize their potential in non-academic cognitive areas, such as the arts. Whether or not such non-cognitive outcomes can be connected to effective student-centered instruction is, however, a subject of another research project, and possibly more than one. Meanwhile, the true lack of, or a negligible presence of an association with math and science student test scores in the population of students across Kazakhstan and other post-Soviet states remains a possible explanation of the results I found in this dissertation.

In Chapter 7, I begin with these results as a starting point for the exploration of student-centered instruction as a phenomenon in the post-Soviet state, with an in-depth look at the specific challenges and opportunities facing the education sector in Kazakhstan. I further explore the presence of instructional reform in official state policy on education, and examine how the vision of a "new way to teach" meshes with the realities of state-driven incentives and accountability systems for schools nationwide.

CHAPTER 7.
LACK OF CAPACITY OR RATIONAL CHOICE?

7.1. INTRODUCTION

As I pointed out in Chapter 2 and Chapter 4, student-centered instruction in Kazakhstan occupies a prominent place in policy talk. In 2005, instructional changes were emphasized as a priority area in the national education strategy – presented in the National Education Development Program (NEDP) for the years 2005-2010, put out by the Ministry of Education and Science. That policy document states that "changing the mindset of the teacher, and… shifting the role of the student from a passive recipient to an active participant of the learning process" is crucial for developing the core competencies required for the nation's competitive development in the 21st century. In 2008, President Nazarbayev, the nation's leader since 1990, put educating students to develop "an ability not only to consume, but to create new knowledge and innovation" as the number one goal in his vision statement, Intellectual Nation 2020. These priorities continued in the 2011-2020 NEDP, albeit with a lesser emphasis on "changing the mindset" towards the instructional process per se. The new policy document contends that the main challenge of the new era in the country's educational development is to reorient teaching so that it develops new key competencies such as critical thinking, creativity, and versatility in the workplace, shifting from traditional consumption of knowledge flowing from the teacher to the student.

However, the results of the quantitative analysis described in Chapter 6 show no evidence that student-centered instruction had any impact on the student achievement outcomes in fourth grade mathematics and science, measured by the TIMSS international student achievement study in 2007. To be precise, the coefficients on the variables

measuring student-centered methods were small, and error variance surrounding the estimates was too great to conclusively rule out the null hypothesis of no effect of student-centered methods of instruction on achievement in the population of fourth graders in Kazakhstan. Even without reference to statistical significance and generalizability of sample estimates to the population, the point estimates obtained on proxy variables measuring teaching practices that fall within the definition of student-centered instruction were too small to be of substantive interest: the coefficients, when properly controlled for relevant background variables, ranged from .1 to .12 of a standard deviation of achievement outcomes associated with receiving more exposure to student-centered instruction. In the broader context of all post-Soviet states that took part in the TIMSS 2007 assessment, student-centered instruction also did not show a coherent trend that would indicate its potential contribution to the overall levels of achievement in math and science across these countries. Placed next to its neighbors in post-Eurasia, Kazakhstan did not demonstrate large or consistent coefficient on variables measuring teaching practices in general.

This chapter examines the context of the Kazakhstani education system, particularly the dynamics of the state education policy and its implementation as it pertains to instructional reform. I explore the factors that may help explain why student-centered instruction, having been placed as a high priority on the nation's education development agenda, fails to show conclusive evidence of impact on achievement in mathematics and science. I look at the context of the schools, and the realities facing the teachers and school principals, their views and perceptions regarding effective instruction in general and regarding student-centered instruction, in particular. I also look at the priorities of the government in educational development (stated and unstated), and the incentives and accountability mechanisms created for increasing the quality of education.

7.2. DATA

The data for this chapter come mostly from interviews with educators in Kazakhstan, as well as policy documents and reports on the status of education in the country, which provided the background for the discussions with key informants and focus groups. Chapter 5 (Methodology) provides detail on the interviewees for this study. Because the qualitative data collection was carried out solely as an extension of quantitative analysis, I did not construct a nationally representative sample of educators, but chose a purposeful sampling strategy, where several educators in the area of mathematics and science instruction were contacted first, and were asked to provide the contacts of other knowledgeable peers. At the same time, I sought to ensure that administrative and policy making levels of the education system were also represented in the sample of interviewees. For this reason, the group of key informants includes not only teachers and principals of public and private schools, but also NGO representatives, an official from the Ministry of Education, instructors at the state in-service and pre-service teacher training institutes, and an official from the National Testing Center, which implements high-stakes testing such as mandatory school leaving exams, as well as intermediate assessments in upper middle school. The interviews followed a semi-structured format, where a set of core questions related to the implementation of student-centered instruction were asked of all interviewees, but space was left open for additional views and reflections of interviewees on the changes in state educational policy, the strengths and weaknesses of its education system, and the challenges and opportunities facing Kazakhstan's schools in the coming decades.

As I noted in Chapter 5 (Methodology), it is important to acknowledge my own connection with the data, as a native of Kazakhstan, and a former program officer of a donor

agency, overseeing programs in education in Central Asia. While none of the projects I oversaw were directly engaging with counterparts in Kazakhstan, I was well aware of the context of reform well before beginning this study – which, incidentally, spurred my interest in further research. I also had some ties to the education sector in Kazakhstan, through professional and personal networks. While these circumstances are likely to have opened more doors and helped put some interviewees at ease, a slight possibility cannot be ruled out that the findings may have been influenced by the respondents' reaction to my background.

Interviewees spoke on the condition of anonymity, which was particularly important for teachers and school principals. Before making statements that they perceived as potentially diverging from the official rhetoric or language of the policy documents, teachers and principals requested additional reassurance that no part of their opinions will be explicitly linked to their identity in this study. Teachers seemed wary of potential repercussions that might come from the state governing apparatus, should an impression be formed that they disagreed with the position of the state, or that they were pointing out the weaknesses of the system to a greater extent than its strengths. Education officials at higher levels did not request reassurance of complete anonymity, but emphasized that their opinions were exclusively their own and did not necessarily reflect the official views of their agencies as a whole.

7.3. ANALYSIS: STUDENT-CENTERED INSTRUCTION IN THE CONTEXT OF THE SCHOOL

As a first step in building an understanding of the post-Soviet education system as an interpretive context for my quantitative findings, I focus on the level of the school and the teachers, exploring their perceptions, attitudes, challenges, and opportunities which

have a direct bearing on the choices they make in their instructional approaches. These qualitative findings follow several key hypotheses for the lack of a relationship between student-centered instruction and achievement on TIMSS. First, a negative attitude towards student-centered methods on the part of the teachers, or an uneven interest and implementation across different schools may predict the lack of an association with achievement, or, worse yet – a negative association between engagement in student-centered methods and student achievement outcomes. Secondly, a lack of familiarity or training in instructional methodology may also make it difficult to expect a robust positive association. Finally, it is important to understand both the amount of open space for teachers to experiment in new methods, as well as the external incentives that determine their choice of instructional methods and classroom practices.

7.3.1. PERCEPTIONS AND ATTITUDES TO STUDENT-CENTERED INSTRUCTION

As I suggested in Chapter 6, one of the potential explanations for inconclusive results of examining the relationship between student-centered instruction and student achievement could be a lack of coherence and consistency of implementation of such teaching practices in the classroom. If there was no understanding of the learning theory behind student-centered methods among teachers, or if the understanding of theory varied according to the personal perceptions of teachers, one would be hard-pressed to find a consistent approach to the use of student-centered methods in any subject, let alone the more technical ones, such as mathematics and science. Furthermore, if teachers did not believe that student-centered pedagogy could bring positive results to learning, or if they were convinced that traditional teacher-centered "rote" methods were more effective, student-centered instruction would again be unlikely to take root in mainstream teaching,

and hence, unlikely to show a visible association with higher achievement in these areas beyond a few isolated cases. Therefore, as part of the qualitative case study, I asked educators to share their views on effective instruction and their attitudes to student-centered methods vs. traditional approaches.

7.3.1.1. ATTITUDES

Most interviewed teachers and principals, both key informants and focus group participants, said that they wholeheartedly supported changing the teaching approaches in their subject areas. Teachers in particular were vocal in emphasizing the fact that the students they now teach are "not the same as before", and required substantially greater attention and focus on the part of the teacher on making learning relevant to their realities, and greater effort in engaging them in active forms of learning, rather than relying on lecture-style presentations and passive completion and grading of homework assignments. Students "nowadays" were said to have greater curiosity and possess lower barriers in challenging their teachers by asking additional questions and requesting additional information, and were reportedly showing greater interest and motivation in their studies when . In addition, teachers in the more affluent schools felt the pressure to remain on the same level with their students in terms of their knowledge and understanding of diverse phenomena, based on their comfort and ease with internet-based technology. According to the interviews, many teachers believed that traditional instruction, where the teacher is the source of all knowledge, is outdated and a different approach to instruction was more relevant for the needs and interests of students.

> *Any teacher should use the project-based* [teaching method]*, because it's good for the development of the personality of the child, development of their skills. It's very good for the kids.* (Teacher and school administrator, Focus group #2, Astana)

> *We have students here who write poetry, perform artistically. It should all be targeted at the needs of the child. …There is a shift towards so-called interactive methods. That means it's not that the teacher lectures the content and the student takes notes, but the teacher poses a task, a problem, and the teacher and the student solve it together during the lesson. That is called individualized instruction.* (School zavuch – Assistant Principal, Almaty).

While student-centered instruction seemed to be a favored and widely accepted as beneficial for learning in the modern-day classroom – at least verbally if not in the actual practice of teaching – a few dissenting voices were also present among senior teachers. They argued that while student-centered methods such as group work, project based learning, and use of graphic organizers was attractive and exciting for students, there was little to be gained from them in terms of actual results in mastering subject matter that the students can visibly demonstrate at an assessment. They also argued that active learning approaches were placing an unfair share of burden for finding the knowledge and properly understanding it on the student, instead of holding it firm in the hands of the teacher. Smarter students might be successful in pursuing a project and learning the subject matter correctly at the same time, said these respondents; yet a weak student is unlikely to draw the right conclusions if he or she is not explicitly told what is important and what is to be learned from a given piece of instructional material, by the teacher.

> *I believe the teacher must teach – must transfer her knowledge to the student. There are new experimental methods now, where the students are asked to learn things on their own. But the kids are all different: there may be one who reads the textbook once and understands it, and another* [kid] *who reads it ten times and still doesn't get it. So it's the job of the teacher to teach, and for that we need highly qualified teachers.* [Teacher of mathematics, focus group #1, Almaty]

Thus, some teachers were wary of shifting the active role in the instructional process from the teacher to the student, which they perceived to be an unjustifiable risk to the effectiveness and completeness of learning their subject matter. Individual assignments where students work on their problem sets, which are then corrected by teachers and in

many cases, discussed with the class as a group, provided, in their view, a "tried and true", and proven pathway for improving achievement outcomes on student exams – be it formative assessments during the school year, or the high-stakes exams at the end of compulsory school program.

7.3.1.2. LEVEL OF UNDERSTANDING OF STUDENT-CENTERED INSTRUCTION

Regardless of their attitudes to student-centered instruction, the teachers found it difficult to describe the learning theory behind it, and were unaware of evidence that would support of a positive attitude to such a teaching approach. Few teachers displayed any knowledge of the theoretical underpinnings of student-centered instruction, and their responses were often based on their own experiences in the classroom, or the experiences of other teachers that they had heard about, rather than on formal training or the reading of published sources on child-centered methods. When asked to name the characteristics of an effective teacher, the responses included the ability to engage the students and put them to work in seeking out knowledge themselves, to build instruction around the individual learning styles of the students, and to help them to retain what they learn over the long term. Almost all responses also emphasized the need to be versatile with educational technology. Only two of the interviewed teachers brought up the names of Russian educators who reportedly had published on non-traditional types of instruction, and only did so in passing, without providing an overview of what these non-traditional methods entailed.

When asked about their freedom to choose an instructional method they deemed most effective, all interviewed teachers reported that they were not limited to any particular instructional methods, provided they fulfill the requirements of the State

Standard on Education, which stipulates the content and numbe of instructional hours to be devoted to each piece of the curriculum. However, this freedom often did not translate into the choice to try non-traditional approaches, as teachers were often unaware of where to start with student-centered instruction.

> *There are many different methodologies that are practiced in schools these days... But there isn't a systematic way in the choices of instructional methodology. Every school, every teacher does as they please... And because of that, overall quality of instruction suffers. But there should be some common knowledge, some shared understanding across the education sector of what methodologies exist, and their benefits and disadvantages, their outcomes.* [Former school principal and head of a private test preparation center, Almaty,]

The lack of clear guidance and leadership on the part of state education officials in the aspect of instructional methodology resulted in confusion among teachers over which methods are considered better, more effective, or more up-to-date. Teachers also lamented the lack of clear guidance on how to practice student-centered methods, and what specific requirements there were on their lesson planning, if, in fact, student-centered teaching is the approach preferred by the state education authorities and endorsed by the President. In general, there appeared to be a void in the implementation capacity of the state education authorities concerning student-centered instruction, and teachers were unsure if they were to fill this void with methods of their preference, or follow the practices established in the Soviet period. As a result, this void was often filled by approaches or strategies with which the teachers were most comfortable, and at levels that the teachers themselves were able to attain through their own personal experiences. Even younger teachers found it difficult to function without coherent guidance that the schools – particularly the senior teachers and administrators with substantial experience in the Soviet system – have come to expect from the central educational authorities. They reminisced about the Soviet-era strict guidelines

regulating the teaching process to the minute, which they believed were not replaced by equally comprehensive guidance:

> *In the past, we had everything regimented, everything was clear, which methods we are using when. Right now – and maybe these are still growing pains – but we don't have a clear, consistent approach to teaching. The new teachers, they may be ready to be creative, but there is nothing, no foundation on which to experiment. Everyone is on their own, and they do as they like in terms of instructional approaches. [Teacher of mathematics, Astana, Focus group #2]*

As this latter quote points out, there was recognition among educators of the desire to be innovative and creative in teaching, especially on the part of newer teachers, who subscribe fully to the rhetoric of "change in the mindset" of the educator. However, without proper training and clear direction from the state, teachers were unsure on which core framework to rely, and uncertain of whether their efforts were likely to improve student achievement outcomes, or would simply cause confusion. Some were openly skeptical about the possibility of a high-level strategy to reach the level of the classroom, due to the lack of support from the state for the teachers, and saw a disconnect between the rhetoric of the strategy and policy documents and the realities facing teachers:

> *Well, the* [education development] *strategy is all good, is very well developed.* [President] *Nazarbayev talks all the time about the importance of supporting teachers. But… the implementation – it always hinges on the human factor, on the management, administration. That plays a huge role in education. But as for the strategy – yes, it's very good.* (School administrator, Almaty school #134)

Summarizing the overall attitudes towards student-centered instruction, it is fair to say that student-centered methods generate supportive rhetoric not only in policy document and overall vision statements, but also in discussions among teachers. However, when it comes to implementation, teachers are uncertain of what learning theory lies behind such methods, what outcomes can be expected of them, and – most importantly – how to build a coherent, comprehensive, and effective instructional environment based on

student-centered methods in their classrooms. Two decades after the breakup of the Soviet Union, the teachers continued to rely on the state to provide guidance and direction in setting instructional priorities and carrying out the practices that were seen by the government (and not always by teachers) as more relevant to the needs of the current generation of students.

7.3.2. QUALITY OF TEACHER TRAINING

The quality of teacher training – both the initial, pre-service training and professional development – is a crucial ingredient of a competent teacher. While there may arguably be intrinsic qualities that distinguish especially effective or charismatic teachers from their less successful peers, one can hardly expect a poorly trained teacher to effectively deliver curricular content at any level of complexity to her students. In my quantitative analysis, I included the levels of formal education attained by the teachers, along with variables indicating their specialization in mathematics or science, as covariates in models measuring the effectiveness of student-centered instruction. The results did not show a substantively significant association between higher level training or specialization and student test scores, which may be explained by the lack of sufficient variation in the level of training or the content of specialized vs. general education programs among teachers. Through interviews with teachers, principals, and teacher training instructors, I gathered qualitative information on the content of training programs, and the availability of training in student-centered methodologies.

7.3.2.1. PRE-SERVICE TRAINING

The lack of confidence in student-centered methodology on the part of experienced teachers might have been remedied by an influx of new teachers trained in the teaching

approaches that the state pledged to promote throughout the national education system. However, the quality of training in pre-service institutes in Kazakhstan is reportedly one of the weakest links in the system, both in terms of the content of training and in the proportions of pre-service graduates who actually intend to enter teaching as a career. Training in methodology is said to be lacking in general among the new graduates of teacher training programs. Respondents of this study were almost unanimous in their criticism of the level of newbie teachers' versatility with instructional methodology. While all respondents lamented the fact that teaching as a profession fails to attract talented youth (due to low entry wages), they also commented on the lack of quality instruction in the standard pre-service teacher training programs at state-run universities:

> *I see young teachers coming into the school, without any knowledge of new methods. They get trained in old, traditional methods of instruction, whereas the schools now search for new ways. So you have to spend time and effort on making them learn how to use new techniques.* (Interview, official of the Ministry of Education and former school principal, Astana)

> *There is emphasis on content in teacher training programs, but not on methodology. And what is a teacher who doesn't know methods* [of instruction]*? A self-trained workman.* (Interview, in-service training instructor, Almaty City IPK)

> *I have hired a few of the recent graduates - they are so unprepared, you have to spend the first year training them to simply be able to stand in front of a class, to teach a lesson.., they are scared of the kids... Their subject matter training is very poor; and their instructional competency is absent altogether. For teaching practice, they get one or two months inside a classroom, over their whole five years in college - that's all. They sit in and observe teachers, and the smart ones may pick up some things, but most of them remain at the same level. And then they graduate and they bring their diplomas to us in schools, with zero knowledge. Meanwhile the school is facing new challenges - we need new ways, new mindsets. Where will that come from? Nowhere.* (Former school principal; for profit test prep center director, Almaty)

The National Program on Educational Development 2011-2020 does not include specific provisions for the improvement of teacher training programs at state universities and colleges. An interview with an instructor in a child-centered methodology called Step-by-Step (developed by the Open Society Institute of the Soros Foundation), at a large state

university-based pre-service teacher training program in Almaty revealed a lack of institutionalization of student-centered methodology in the teacher training curriculum. While the respondent herself taught the course on student-centered instruction (notably, in the form of lectures), she did so on her own initiative, based on the training and experience she gained as a Step-by-Step staff trainer in the late 1990's.

> *When we first started, we had 13 colleges and 11 IPK* [in-service institutes] *working with us on this methodology for pre-school. There were special courses for Step-by-Step in several pre-service training institutes. I personally still teach this course. I don't know about other pre-service training institutes. They may have canceled them all because there was no more support* [from the NGOs]. *But I still do it, because it is my area, and I contributed to the development of the state standards for preschool education.* (Interview, pre-service teacher training college instructor, Almaty).

The latter quote refers mainly to the training of preschool teachers, and to lesser extent, the teachers of early grades of comprehensive schools. Early efforts in the dissemination of child-centered methodology were well funded by donors such as USAID, UNICEF, and the European Commission, according to two respondents who were involved in delivering the training courses. With time, the funding from these international bodies dried up, while the expected institutionalization in pre-service training colleges did not take firm hold, according to the interviews. With diminishing funding from international development agencies, and from the Open Society Institute, leadership capacity in continuing the initiative also waned, unsupported by the government structures and regulatory frameworks. However, at the time of the data collection, the government reportedly adopted a new State Standard containing curricular parameters for the training of preschool teachers, and one of the interviewees claimed having contributed a requirement to deliver at least two hours of coursework in student-centered teaching. It remains to be seen, however, what those two hours will cover in reality, and whether this

will be sufficient exposure for teachers to obtain an understanding of this approach to classroom instruction.

7.3.2.2. IN-SERVICE TRAINING

With respect to the professional development of existing teachers, the state in-service teacher training institutes (IPK) have reportedly increased their role in providing the schools with exposure to recent developments in instructional methodology. Schools in Almaty reported working regularly with the city IPK and using them as their source of instructional support; and the respondent from the Almaty IPK said the timetable of in-service training workshops is meticulously followed, with full sessions several times per year, and shorter seminars in the interim. It is not clear, however, what proportion of the in-service training courses have integrated some form of student-centered, critical thinking approaches, and what proportion remained traditional lectures. One training session observed for this study at the Almaty City IPK – a course for upper-grade chemistry teachers – was purely lecture-based, and focused on refreshing subject matter knowledge of some aspects of chemistry, rather than improving the methods of instruction. Similarly to the situation at pre-service teacher training programs, professional development courses may or may not include training in methods, depending on the personal interest and competence of a given IPK trainer. Further, the inclusion of content on student-centered instruction depends entirely on the instructor's attitude towards the approach and his/her familiarity with the largely NGO-driven methodology packages such as Step-by-Step or Reading and Writing for Critical Thinking (RWCT).

In-service training institutes from areas outside of the two main cities reportedly have gone further towards incorporating courses in student-centered instruction into their regular curriculum. According to the interview with the director of the local Step-by-Step

Foundation in Almaty –one of the largest and oldest education nonprofits in Kazakhstan – originally there was genuine interest in the content and methodology of child-centered instruction for preschool and early grades in regional in-service training institutes and local departments of education. In recent years, professional development courses based on Step-by-Step content, according to the respondent, were adopted in several preschools in Karaganda, Pavlodar, and Atyrau regions. Schools and parents reportedly showed great interest and support for child-centered methodology in preschool and early grades, "because they immediately see the results in their children" (interview, regional education NGO representative, Bishkek). However, support was less prevalent in middle and upper grades, as more emphasis was placed on test-based accountability, on which the student-centered instruction did not appear to be making visible impact.

Full completion of training also does not always guarantee sustained implementation by teachers, due to the challenges and disincentives described above. As one teacher, certified in Reading and Writing for Critical Thinking (RWCT), a methodology for upper grades, shared:

> *Yes, I passed a full course of RWCT certification in 2002. I don't do it as much anymore... But if there is an opportunity – and that does not happen often – to show what we can do, like in an "open lesson" or something like that, then of course we like to apply these techniques, and the kids appreciate that too, it's fun for them.* (Teacher, focus group #2, comprehensive general school, Astana)

As this quote demonstrates, with one-off training events in effective instructional methodologies such as RWCT, but no support, follow-up, or real incentives from within the system to keep up the use of newly adopted methods, teachers lose momentum and revert to their comfort zone, applying the usual practices in which they had been trained in their formal education, or to which they got accustomed to during their careers as teachers in mainstream schools.

7.3.3. CLASSROOM IMPLEMENTATION CHALLENGES

7.3.3.1. TEACHING OVERLOAD

While guidance and support of instructional practice on the part of the state was, according to some interviewees, a crucial missing link between the rhetoric of instructional improvement and its implementation, it was not the only one. Most teachers – with the exception of teachers at a semi-private elite school in Astana, said they were overloaded with work and therefore unable to take on the extra effort required to go beyond traditional lecturing and questioning. According to teacher interviews, they found it hard if not impossible to make a decent living on a single *stavka* (a full-time teaching load), currently set at 18 hours of class per week, with a starting salary of KZT 24 thousand per month, or approximately $200. For this reason, most teachers signed up to work two or even three *stavkas* to increase their take-home pay, in addition to the private tutoring jobs that senior and more experienced teachers had on the side (on this topic see UNICEF 2011, forthcoming). It was difficult to obtain a precise number of how many hours, on average, an average teacher spent in the classroom, but the estimates given by interviewees ranged from 32 to 40 hours of actual class time, leaving almost no time for out-of-classroom work, such as perfecting the methodology for one's lessons, seeking out new materials to use in class, or engaging in professional development. When asked whether they agreed with the statement in the 2005-2010 National Education Development Program that the there was a need for a change in the teachers' mindset, so that the pupil becomes an active participant in learning, rather than recipient of knowledge, a teacher and *zavuch* at a school in Astana responded:

> *When will they have time for the extra stuff? They are barely making it with the required curriculum. They come home and collapse… well, actually, because they are mostly women, they still have their house chores to take care of… So these good*

> *aspirations of the national policy makers, to change the mindset of the teacher, those are all good intentions, but they will only come true when the teacher starts to feel like she is a human being, and not like a tired racing horse.* (Interview, comprehensive general school, Astana)

These findings reflect the fact that despite the greater level of overall wealth in Kazakhstan, compared to its neighbors in the region, the practice of working multiple loads may be as prevalent among Kazakhstan's teachers as in the other states, thereby concealing the actual magnitude of teacher shortages. The challenges of the stavka structure in the post-Soviet education systems reflect on the status and the capacity of the teachers, and some observers argue that structural changes are necessary for any quality improvement to take root (Steiner-Khamsi, 2007; UNICEF forthcoming, 2011).

While the overload of teaching hours as a result of multiple stavkas per teacher is troublesome, it is certainly not a new phenomenon. Working more than one teaching load was normal during the Soviet period, and new teachers were even encouraged to take on a greater number of instructional hours, both to improve their teaching skills and to increase their pay, according to veteran teachers interviewed for this study.

> *In the old days, there were mentors who would tell us: "Sweetheart, do take the extra teaching hours. Don't deny this opportunity, it is good for you. But today, the experienced teachers don't care if the young don't have enough teaching hours to pay their bills. They will never offer advice or mentorship…"* [Retired teacher and current in-service teacher trainer, Almaty].

The flip side of teachers teaching more than one *stavka* was that despite a general shortage of good teachers (noted by all interviewees, as well as in national policy documents), fewer teaching positions were becoming open to newcomers every year, especially in cities. Older and more experienced teachers were taking the extra hours willingly, and were reportedly reluctant to give up their hours as it meant a decrease in pay, according to the Ministry of Education official interviewed for this study (Kusherbayev,

personal communication 2010). Teachers in Almaty admitted that there was a seeming lack of space for young teachers to take positions in urban schools that had weight and prestige, and noted that new graduates of teacher training programs were encouraged by government subsidies to go out into the countryside and teach at rural schools. (Focus group #1, Almaty 2010).

In addition to the instructional burden that comes with taking a double or triple teaching load, all interviewed teachers (with the exception of the aforementioned elite private school in Astana), said that a large proportion of their time was spent filling out reports as routine data collection by the government. The reports focused on a variety of aspects related to instruction and school management: the amount of time planned and actually spent on the delivery of the national curriculum, by topic; attendance and discipline in the classroom; hours spent on peer learning (observing lessons of other teachers); and textbooks used in lessons and for assigning homework. This indicates that the state education management information system in Kazakhstan operates using statistical information reported by teachers on a regular basis. However, while such detailed routine data collection is undoubtedly useful for monitoring purposes, the amount of time it requires to produce, according to the interviewees, eats into the work hours that the teachers might otherwise have spent preparing for class, further reducing the possibility of their putting additional effort into improving the effectiveness of their classroom teaching practices.

In sum, even if instructional guidance were present and state-of-the-art training sessions and materials in student-centered teaching methods were available, effective instruction that requires substantial preparation for lessons and an individualized approach to the students would be difficult to attain for most teachers in mainstream schools, due to

the large burden of classroom hours that they take on in order to increase their pay. No incentives are currently present for teachers to reduce their face time in the classroom in favor of greater effort in preparation, in professional development, or in instructional innovation. In addition, routine paperwork demanded of the teachers is said to diminish their satisfaction from their profession, to eat up what little time there is for pre-class preparation, and to turn away new teachers who are overwhelmed by what they see as duties unrelated to their core responsibilities of teaching their subject matter. It would be difficult to expect that the rhetorical acceptance of "placing the child in the center of the learning process" would bring about any improvement in outcomes given these challenges facing teachers in Kazakhstan.

7.3.3.2. STANDARDIZED TESTING

A growing emphasis on standardized testing as a means of school accountability in Kazakhstan was also cited by teachers and school principals as a disincentive for using open-learning methods, particularly in upper secondary grades. Because no data is available at this time on Kazakhstan's 8th graders' performance on TIMSS, I was unable to triangulate the view of several respondents that the use of student-centered methods in upper grades is almost nonexistent compared to primary school, due to the high pressure on students and their school administrators to show high scores on the nation's unified high-stakes school leaving exam – the ENT. Teachers and school principals shared in the interviews that they are given target ENT scores they must attain in a given academic year, as well as regularly informed of the expectation of year-to-year improvement of the score. The pressure is particularly intense on school principals, district officials and other midlevel administrators, who reportedly may suffer dismissal as a result of not attaining an ENT

target. Upper grade teachers and principals agreed that memorization and repeated drilling in the format of the ENT test are the most proven pathways to higher scores.

> *The way we prepare the kids for the test, we drill them, we make them memorize content; memorize answers to the test items. Questions and answers. But you really need* [the student] *to be able to connect that content logically, build linkages among discrete pieces of information. How would they do that, if they have memorized the questions and answers? But that's how they train for the test. I do it, too, in my prep courses. They memorize 1000 items and get into colleges. At least 40-50% items are those they have already seen before and remembered.* (Interview, former school principal, currently director of test prep center, Almaty)

Arguably, the pressure of standardized high stakes testing hardly helps to explain the lack of a conclusive relationship between student-centered methods in fourth grade math and science. However, it signals a dynamic that focusing on memorization may be quite rewarding, in terms of actual measurable test scores, and consequently, the academic standing of the school and the teacher. In the absence of another incentive that rewards the skills that the government has flagged as "key competencies" – critical thinking, creativity, ability to deconstruct factual information and to apply knowledge to reality – the likelihood that student-centered practices would take root is lower.

> *The teachers right now are caught between two pressures. One is the pressure to deliver content training, prepare the kids for the ENT and other tests. And on the other hand, we have this movement, this fashion to develop modern European trends – the competencies, critical thinking, and ability to apply knowledge.* (School assistant principal, Astana)

A precursor to standardized testing as a measure of overall school quality, academic Olympiads, by contrast, stimulate individualized instruction and the development of strong critical thinking abilities in students. Olympiads gather the most academically able students in each subject, at different levels starting from school and district to national and international competitions. Before the ENT, the number of Olympiad winners among one's students was seen as a matter of prestige and status, and taken into account for teacher

promotions and pay increases, thus acting as a powerful incentive to engage in individualized instruction. However, in most schools the Olympiad participants consisted of less than 1% of all students, and were often given specialized preparation outside of normal class hours. Therefore, while the Olympiads remain a measure of success alongside ENT, they have little impact on the mainstream instructional approach.

In sum, the realities of the teachers in Kazakhstan – the confusion over what constitutes effective instruction in terms of developing the key competencies outlined in the policy documents; the lack of training, support, and guidance from the state apparatus that teachers rely on, and haphazard interest in student-centered methods – offer some insight into the lack of relationship between these instructional practices and achievement found in the TIMSS 2007 data. As a next step, I look at the overall dynamics of the state education policy implementation, with the goal of shedding light on the state policy agenda, its level of overlap with official rhetoric, and the limitations and challenges facing the state as the key driving force in educational development in the country. As part of this analysis, I examine the plausibility of the hypothesis that the centralized state has greater mechanisms for direct and unhindered implementation of its development vision (see Chapter 4), and is therefore more capable of aligning its policy rhetoric with action. This portion of the case study on the state also helps me to address the seeming disconnect between the state promotion of student-centered pedagogy as an inherent and necessary part of educational development, and the complete lack of support for its implementation, revealed in the interviews with teachers and school principals.

7.4. STATE LEVEL ANALYSIS: A LAISSEZ-FAIRE APPROACH TO QUALITY

One of the core assumptions behind the conceptual hypothesis, based on Carnoy et al. (2007) of the centralized state, as an entity capable of effectively and efficiently implementing is policy vision, is the presumed tightness of coupling that is necessarily present in a centralized environment. That is, the hypothesis goes that if the post-Soviet state with a strong central authority desires to completely revise its approach to instruction and thereby create a generation of citizens possessing a key set of competencies rewarded in the modern society, that this kind of state has better mechanisms at its disposal, because its hierarchical administrative system easily responds to top-down initiatives and carries out the vision in accordance with the grand plan. To make it more specific, if student-centered instruction is the means for change, then the state ensures that it is implemented throughout the education system.

However, the results of my quantitative analysis showed no conclusive evidence of association between student-centered instruction and achievement in Kazakhstan (albeit, not unlike in other post-Soviet states). The limitations of the TIMSS data notwithstanding, more importantly, interviews with educators revealed that the realities of instruction in Kazakhstan's public schools have little changed from the Soviet era, and while the teachers adopted the language of student-centered methodologies, they had little idea of what learning theory was behind them, or how, when, and why such methods could be used in their lessons, or how to use them effectively in combination with traditional approaches. Implementation of instructional methods is based more on personal interest on the part of the teachers, and confounded by their perceptions of what is interesting to their students, rather than on a coherent strategy and support from the state.

These two core findings seem to suggest that the state, or more specifically, the government of Kazakhstan, is not necessarily in a better position to carry out its vision in instructional innovation as it may seem from its level of centralization, or that centralization does necessarily mean tight coupling between the elements of the system. However, it may also be that while the state declares instructional reform as the core to its educational development, it chooses to attain it with measures other than training and professional development of teachers. Rather than evidence of loose coupling in the education system, the lack of adequate training and instructional support for teachers and the seeming void of leadership on the part of the state appear to signal a more libertarian approach to teacher quality improvement. Recognizing the lack of central capacity to follow through on the measures necessary to carry out a top-down quality improvement strategy, the state embarks on large-scale, visible improvements in the school infrastructure, while letting teacher quality be induced through market-driven policies such as vouchers for professional development and tighter restrictions on entry into the profession and retention of the teaching license. Furthermore, the state shuns the involvement of third parties in its education system, despite their acceptance by teachers, and focuses on elite-building projects such as the establishment of gifted schools and introduction of system-wide ability tracking policies for mainstream schools.

7.4.1. LIBERTARIANISM IN TEACHER QUALITY IMPROVEMENT

Despite the loud and clear message in the two National Education Development Programs (2005-2010 and 2011-2020) and the rhetoric of the President about the need to reform in mindset of the teacher, no concrete plan of action was outlined to build a training system that would train teachers of the new generation. State pre-service and in-service training institutes continue to rely on the initiative and personal interest of their staff in

offering the training in instructional methodology. In contrast to the study that served as the foundation for my conceptual framework, none of the factors that Carnoy et al. (2007) described as crucial ingredients in Cuba's academic success – the tight linkages between the curriculum developers, pre-service and in-service training colleges, and schools, as well as the coherence of all elements of the system, with strong capacity at the center continuously supporting local instructional leadership – were present in Kazakhstan.

At the same time, none of the interviewees of this dissertation doubted the state's commitment to education as an overarching objective. All educators had noted the enormous influx of state financial resources into the schools (but not pre-service training institutes). The physical infrastructure of the schools was being renovated, teacher salaries were being raised, and new technology, including new science equipment, was firmly entering the daily routine of urban schools. The level of domestic financing for education tripled in absolute terms, and grew from 3.4% GDP to 4.0% GDP between the years 2004 and 2008, according to official figures from the Ministry of Education (MOE 2009 Annual report). The average teacher salary in the general comprehensive schools across Kazakhstan was raised to KZT 27,940, or approximately $230 (exchange rate source: National Bank of Kazakhstan archives) in 2008, up from KZT 15,000 or $100 in 2004, per one teaching load (*stavka*). The Ministry of Education reported that this number put the average teacher salary at the 62nd percentile of wages across all sectors of the economy – without taking into account the fact that most teachers worked more than a single stavka, according to those interviewed for this study. School principals in Almaty and Astana said that increases in teacher salaries were done on an annual basis.

The provisions of the 2011-2020 NEDP put forth the argument that in order to increase teacher quality, one has to improve the selection and retention of individuals who

are *intrinsically* better able to teach effectively. The strategy document seeks to "raise the status of the teacher" by raising salaries, restricting admission into teacher colleges, re-establishing additional teacher certification procedures, and introducing teacher choice of the providers of in-service training, through a pre-paid voucher mechanism. Teacher working conditions are to be improved through school renovations and the provision new equipment and supplies to outfit classrooms, including lab equipment and new technology, such as "interactive blackboards" – classroom board-size screens connected to the teacher's desktop, with pre-loaded educational resources and access to the internet. While no language is included in the NEDP 2011-2020 on reforming or strengthening the capacity of teacher training institutes, the policy document has a provision for teachers to receive professional development vouchers, which they would be free to use with a training provider of their choosing. According to teachers and NGO professionals, the market for such services is far from plentiful, however, and it remains to be seen whether the choice offered to the teachers results in higher quality professional development they receive.

At the same time as increasing teacher salaries and improving working conditions, as I noted above, the NEDP 2011-2020 includes provisions that require stricter selection into pre-service teacher training colleges, such as personality testing designed to screen out candidates psychologically not fit to be good teachers (MOE 2010). In addition, the policy document calls for a regular re-certification of teachers, based on standardized tests. The tests are designed to measure the teachers' knowledge of their subject matter (50% of test items), understanding of child psychology and pedagogy (25%), and knowledge of "legislation of the Republic of Kazakhstan" (25%), which includes the Constitution, the Law on Education, and the Law on the Rights of the Child (MOE, 2010). Given the fact that most if not all of the teacher cadre were trained in methods and content incompatible with the newly declared emphases in instruction, these policies seem to encourage teachers to seek

out ways and opportunities to build their competency in their field, with the state acting as a licensor rather than the agent of teacher quality improvement.

The overall premise behind the efforts of the government in teacher quality improvement is that by raising teacher salaries, improving working conditions, restricting access to pre-service training programs, as well as imposing greater certification requirements for experienced teachers, the state education system would gradually attract higher-caliber individuals into the profession, and screen out individuals who do not possess the intrinsic qualities required to be a good teacher. This approach to recruitment and retention, coupled with shifting the burden of finding the appropriate professional development opportunities on the teachers themselves would alleviate the need to rely on the capacity of state institutions to carry out quality training and provide instructional materials to teachers. In other words, external pressures are applied to stimulate internal quality improvement. Given the dearth of capacity to provide continuous leadership and instructional support to teachers from the center, this libertarian approach may be a rational way for the state to demonstrate its resolve to raise the bar on teacher quality, without expending the effort necessary to train average individuals to become effective teachers.

7.4.2. RELUCTANCE TO ADOPT NGO-DRIVEN MODELS

While the Ministry of Education and it various agencies are responsible for delivering every element in the provision of education in the Republic of Kazakhstan, NGOs have played an important role in cultivating interest in instructional reform. Open Society Institute of the Soros Foundation, a Budapest-based NGO with a special focus on Eastern Europe and the former Soviet Union, is by far the oldest and for years the biggest actor in

grassroots instructional change. In the early 1990's, the organization began to engage with schools in the region, offering training programs in student-centered instruction: Step-by-Step in early childhood and primary school levels, and Reading and Writing for Critical Thinking for secondary school. In Kazakhstan, the methodologies were introduced in 1995, and remain active in primary schools where teachers went through a standard-length training course, according to interviews with current and former staff of Step-by-Step.

Interviews with the head of the Step-by-Step foundation in Kazakhstan, as well as with teachers and school principals showed that the methodologies were well received by schools, and their use was often a matter of school or teacher professional pride and prestige. As of 2010, some of the regional in-service training institutes (the IPK's in Karaganda, Kyzylorda, Atyrau, Kostanai) had adopted Step-by-Step as part of their professional development curriculum for teachers, which serves as evidence of institutionalization of these programs at selected sites. However, system-wide scale-up and dissemination fell short due to a lack of interest on behalf of the state education policy makers. A regional representative of Open Society Institute's education program that works with schools, providing training in the application of student-centered instructional methods reported that in comparison with neighboring states, Kazakhstan seemed less open to innovation in pedagogy and cooperation with practitioners:

> *… In Kazakhstan, I talk to the Ministry of Education, and I feel like there is no willingness to make any kind of connection. They are really invested in telling me that in Kazakhstan everything is okay, everything is quite great, they know exactly what they are doing, and they have no need of support or assistance of any kind.* (Interview, OSI regional program manager for Central Asia)

At the same time, another program manager, in Kazakhstan, believed that there was increased interest in general quality improvement on the part of the government in the past two years, since the year 2008, and brought up the participation of the Astana city

department of education in Step-by-Step seminars as evidence of this growing interest. Prior to this revival of interest, there seemed to be apathy and inertia in the education system, "there were programs and strategies, but the school and the teacher were essentially left on their own…" (Step-by-Step Kazakhstan program manager, Almaty). The respondent emphasized, however, that a lot of the interest in and adoption of Step-by-Step methods in some regional education bodies rests on the personal interests of individuals in leadership positions. Once such individuals retire or otherwise leave their posts, the same level of engagement is often difficult to keep up. While regional departments are said to be more stable than the highest levels of the Ministry of Education, changes in the center sometimes do ripple through the system, causing reshuffling and as a result, the need to re-establish relationships between NGO representatives and education officials.

Overall, interviews with school teachers offered no evidence of a concerted effort on the part of the state to disseminate the knowledge of student-centered practices throughout the system, but suggested that if teachers found the space in their schedules for learning and practicing the new ways of instruction, they would be praised and perhaps even rewarded. At the same time, the general reluctance of the state to accept the resources and assistance from third parties such as NGOs translated in a lack of willingness to expand on the pockets of innovation that were planted in some primary and secondary schools by OSI and its methodology. Due to lack of an alternative for instructional reform, teachers were faced with the choice of either relying on the rote learning methods, in which they were trained and were comfortable with, or trying a new methodology that seemed to be "in fashion" and risking a loss of instructional quality.

7.4.3. EMPHASIS ON GIFTED EDUCATION

While the quality of education at the national level may be improved by investing resources and technical capacity across all schools in equal measures, and by ensuring both that all students have equal access to the resources of the state and that all teachers follow the officially prescribed curriculum and instructional methodology, it is not the only pathway to achieving higher average outcomes. One may choose instead to focus one's efforts on the weakest elements in the system, such as rural and remote schools – which, for this reason, often become the target areas of international donor-sponsored development programs. Another alternative is to focus on efficiency, providing higher levels of resources to a select group, and a lower level to the rest, with the expectation that the select group would pull up the national averages in achievement studies, and eventually provide the human capacity to push forward the country's educational development. This latter approach appears to be present in Kazakhstan: gifted education, in the form of early grouping of students by ability, is one of the prominent features of Kazakhstan's national education system. As the Ministry of Education put it in the 2009 National Annual Report,

> *The formation of the intellectual elite of the country – of young people capable of taking key positions in the country's governance structures, its economy, science, culture, and art – is one of the priorities in the development of the national model of education (p. 58).*

According to the Ministry of Education (2009), the number of students enrolled in specialized and gifted schools – schools that have selective admission policies – grew almost threefold from around 0.7% in 2004 to over 2% in 2008. While specialized programs, such as schools with in-depth language instruction, or with math and science emphasis, had existed in Kazakhstan throughout the Soviet period, the number of such schools relative to the number of regular schools has grown substantially. This growth can in large part be explained by financial incentives: specialized and gifted schools receive additional funding

from the state to help them implement their new and improved curriculum, which has more instructional hours than that of a regular comprehensive school. By law, gifted and specialized schools must fulfill the requirements of the national curriculum, but can expand and enrich the curriculum in their areas of interest and focus. Specialized schools in Kazakhstan, called *gimnazias* or lyceums, generally select one or two curricular areas of focus – such as languages, math and science, or the humanities. Some schools offer *gimnazia* tracks within a comprehensive school environment, essentially grouping the students by intellectual ability within a single school.

> *Interviews with educators in Kazakhstan revealed a mostly positive view of ability grouping and discipline tracking. The general attitude appeared to be that "not everyone is made for advanced study" and the earlier the system can identify these intrinsic differences and tailor instruction towards educating students according to these differences, the better it is for the students, the teachers, and the nation as a whole.*
>
> *So there may be talented students, but we don't have time to give them their level of problems, because there are weaker students in class, and we have to orient ourselves to the average or weak students. Strong students are left out, because they also require time for higher level problems, and meanwhile we are trying to improve the results of the weak students.* (Focus group #1, Almaty)
>
> *I think grouping by ability should start even earlier [than is envisioned in the new policy document], perhaps around grade 7 or so. This way, we don't waste our time teaching chemistry to children who are never going to need it, but give more to the smart kids who can really do well with intensive instruction* (Interview, teacher of biology and chemistry, Almaty).

The new education policy stipulated in NEDP 2011-2020 embraces this view of efficient resource allocation in education. While at present, all students may choose to remain in comprehensive schools throughout the entire primary and secondary cycles – a total of eleven years, or enter vocational schools after grade 9, according to the new policy this will no longer be a matter of choice for student, but will be decided based on a series of standardized tests. With the addition of the 12th year to the comprehensive (primary and

secondary) school, the policy separates three upper secondary years into academic tracks called “profile schools” (*profilnye shkoli*).

The state has also invested substantial resources into the creation of Nazarbayev Intellectual Schools (NIS) – the semi-private network of schools for gifted students with enriched instruction in math and science disciplines, as well as languages. NIS opened its first school in 2008 in the capital city Astana, and at the time of the research for this dissertation in 2010, the school had been functioning for two full academic years. According to the teachers at the school, and the official NIS website (www.nis.edu.kz), the project is envisioned to expand to twenty schools across the country by the year 2020, in what is seen as the national laboratories of excellence. The highly selective admission of students to enroll in NIS, their location in upscale buildings with the most up-to-date technology, and competitive recruitment of teachers into these schools are all intended to ensure that these are the institutions that will breed the future elite of the country. At the same time, interviews with teachers in Astana indicated that recruitment of NIS teachers fell short of a nationwide search, and selection criteria were less than transparent, resulting in some educators expressing skepticism that any substantive qualitative difference existed between the teachers of regular comprehensive schools and the teachers recruited into the NIS. The teachers of the NIS interviewed for this case study contended that while they may not claim to be the “best in the nation”, the mere fact that were paid well enough to work only a single stavka, and not required to fulfill other administrative tasks or fill out endless reporting forms for the district meant that they were able to spend more time on their teaching, and as a result, achieve better outcomes than teachers in regular schools. One teacher boasted that while in a regular school, she was unable to push more than a couple of students to win places at the Olympiads in her fifteen-year tenure, at the NIS she had five winners in the first year (Focus group #3, NIS, Astana). The NIS teachers, according to their

interviews also were free to pursue professional development opportunities to improve their teaching methods, and were encouraged to try innovative, child-centered, project driven approaches in their lessons (Focus group #3).

While the literature on gifted students suggests that tracking and ability grouping benefits higher achieving students from better-off socioeconomic groups, while detrimentally affecting weaker students, neither this case study nor the brief quantitative analysis of the TIMSS student achievement and school survey data offered evidence of a negative effect for any subgroup of students. In fact, the TIMSS data showed that students studying in schools with ability grouping for mathematics teaching had higher results on the math assessment, and this association is even greater among Kazakh-speaking and poorer students. Schools that track students by ability were also substantially more likely to engage in student-centered methods: 53% of "treated" students (as defined by this dissertation, see Chapter 6) were in schools that group by ability, compared to only 36% of "control" students, controlling for all available proxies of student socioeconomic status and the qualifications of their teachers. However, an array of unobserved characteristics may be affecting this breakdown, particularly as it is not known which students were tracked into high-ability classrooms, and which remained in lower ability groups. More research is required to establish the nature of these relationships, and it is not clear at this point whether investment of more resources in gifted and specialized schools results in greater use of student-centered methodologies.

Based on the interviews and policy documents, it is evident that the development of gifted education, as well as the grouping of students by ability in mainstream schools is not a sporadic process, but a comprehensive strategy by the state to nurture a new elite from the academically higher achieving students. The notions of university education not being

for everyone, and the need to ensure that the best students have the maximum opportunities to learn and contribute to the country's economic competitiveness (NEDP 2010) reflects an explicit desire on the part of the state to maximize efficiency in the use of resources. Given the high correlation of gifted education with urbanicity and, based on interviews, on higher socioeconomic status of the families, such a gain in efficiency may well come at the expense of equity. In primary grades, however, gifted education was less widespread at the time of data collection, although the growth of specialized gimnazias reflected the interest both among educators and the state in building a system where the student's intellectual abilities would be identified and capitalized on early in their academic life. It appears that lacking sufficient capacity to deliver high quality instruction across the whole spectrum of schools at this time, the government uses existing expertise, along with high (and at times, excessive) investments in the infrastructure and technology, to create concentrated environments for intensive education, tailored to students that are most likely to absorb such intensive instruction. Incidentally, such students are also most likely to do well without such extraordinary investments.

7.5. CONCLUSION

In this chapter, I explored the qualitative aspects of the education system in Kazakhstan, examining the realities and challenges that are facing teachers and school administrators, as well as the actions and priorities of the state in building the conditions for continuous quality improvement. I started this study with a hypothesis that in a highly centralized, even authoritarian, post-Soviet system like Kazakhstan, the state would employ the most direct means of achieving quality: by building and disseminating a coherent and comprehensive strategy for instructional improvement, and ensuring its universal implementation through its tightly coupled elements. This hypothesis was driven by the

Carnoy et al. (2007) argument that showed that the most effective pathway to quality was at the disposal of tightly regulated centralized states with the power to restrict individual choices for the benefit of the common narrative. However, in the case of Kazakhstan, instruction in student-centered methods that appeared to be at the forefront of the new narrative for education did not show substantively significant results in the TIMSS 2007 student assessment. The qualitative data offered greater depth to the understanding of this lack of a conclusive relationship, as well as depth to the understanding of the complexities and challenges facing a centralized state with a grand narrative and low capacity, and, it appears, lack of genuine political will to implement it. While the rhetoric of change and innovation was present, and educators were curious about it, there was confusion about what change really means, pride in the old Soviet heritage, and lack of clarity on future directions, with schools given incentives to strengthen the status quo. The government, in its turn, appeared to support student-centered instruction on a discursive level, as an acknowledgement of the global reach of this concept, but without putting real changes into place in the instructional environment in Kazakhstan. To an outside observer, the focus of government efforts would appear undoubtedly the increased resources for equipment and infrastructure, as well as the emphasis on gifted education, and not instructional reform – and especially not one that creates a different teacher-student dynamic in the classroom.

Notably, the case of Kazakhstan presents an example of a deviation from a strong state model of Cuba: here, the state does not hold complete control of the education system, but lets the market play its role – for example, in the distribution of vouchers for training, and in its desire to restrict access to teacher training colleges and raise salaries, with the goal of selecting the best talent. Whether this is the best option for the state and its recognition of the "power of the market" – or simply a gap-mending solution in the absence of state capacity to exercise leadership in instructional reform, is largely a matter of

opinion. The emphasis on the rhetoric of change, however, as well as on competitiveness of the nation as the ultimate goal of the education system, leads the observer to conclude that it is, in large part, a genuine "laissez-faire" approach, and that even provided with the high quality teacher training system, the state would wish to relieve itself of the responsibility to provide all training, and rely on the market to play its role.

Overall, the interviews, focus groups, and document review conducted as part of this study showed that improving the quality of education, as it is measured by assessments of achievement outcomes, is indeed an area of great interest and genuine priority for the state in Kazakhstan – a similarity with the findings of Carnoy et al. (2007) in Cuba. The rhetoric of President Nazarbayev calls for new ways, new skills, new competencies to be developed in the young generation of citizens in Kazakhstan, with the ability to not only effectively consume knowledge, but to generate new ideas and to innovate, as the main challenge of the country's education system. The national education policy documents reiterated this rhetoric, calling for a new shift in the teaching and learning process, making the student an active participant, rather than a passive recipient of knowledge transfer. Furthermore, the government substantially increased the amount of resources flowing into the education sector post 2000, raising salaries for teachers and providing the necessary funding for school maintenance and renovation on a yearly basis, and even greater increases are envisioned during the years 2011-2020, as stated in the draft new NEDP (MOE 2010).

The rhetoric of broad-based instructional reform – the shift towards critical thinking, and an individualized, student-centered teaching approach – does not, however, match the existing performance incentives in the education system. The structure of teacher pay– based on classroom teaching loads (*stavkas*), without a limit on the number of stavkas per teacher – rewards excessive work hours in the classroom and diminishes the

importance of off-class work time for adequate preparation and professional development of teachers, serving as a disincentive to teachers who wish to move away from rote learning methods. In secondary school, standardized assessments, and particularly, the Unified National Test (ENT) – a school leaving and university admission test for students, and the main measure of overall performance for school principals and individual teachers – reward test preparation practices such as the reorientation of instructional time for rote memorization of subject matter content and drilling students in test-taking strategies. Academic Olympiads continued to reward highly individualized instruction to a few strongest students at each school. Given these incentives, the rational response is for teachers to limit student-centered instruction to a few gifted students, and teach heavily "to the test", particularly in the last two grades of secondary in order to show higher average scores. Indeed, interviews with teachers and other educators in Kazakhstan indicate that these practices are commonplace in schools.

There was a substantial influx of state funding into the national education system, with across-the-board salary raises, school renovations, new school construction, and a wide dissemination of educational technology and hi-tech equipment for schools. These inflows created a positive atmosphere in the education system, with all respondents praising the new support of the state and expressing optimism about its development trajectory. However, as centralized and tightly coupled as the state education system was in Kazakhstan, and despite the sharp and visible increase in the level of resources provided to schools, the core element of the educational process – instruction – remained entirely at the whim of individual choices of teachers, dependent on immeasurable and unpredictable elements such as teacher charisma, personal interest in professional growth and enthusiasm. The strategy for the improvement of quality in education, voiced in the two NEDP's, consisted of libertarian strategies such as higher starting salaries, continuous

certification and screening out of lower-ability teachers, and favorable working conditions. None of the policy documents contained a plan for instructional reform. The lack of guidance and the absence of a common understanding of what instructional practices exist, how they can be used in the classroom, and what results one can expect from a given method, created a void in the classrooms that teachers filled based on their prior training, their interests, and their superficial exposure to discussions of teaching innovation.

In interviews, teachers lamented the uncertainty, and those with experience in the Soviet period reminisced about the strict regimentation and synchronization of curriculum frameworks, teacher guides, and textbooks. In the centralized, post-Soviet environment with an authoritarian leader, the state is still expected to act as the provider of quality in education, with a coherent, system-wide approach to instruction and quality assessment. However, the vision behind the national education policy – the shift towards teaching practices that will produce a highly skilled workforce that not only possesses key skills and competencies, but is able to think critically and innovate – was not, at the time of the data collection, directly supported by the state, but hinged on the level of resources in general provided to the schools, with the expectation that instructional quality will follow naturally an increase in funding. At the same time, the expansion of gifted education indicated that efficiency perhaps outweighed equity as a priority in national education policy making. The growth of elite gifted schools and the institutionalization of tracking in upper secondary school serve to diversify educational outcomes and create a heterogeneous social structure as the new graduates enter the labor market. Social capital was thus generated by the state in highly concentrated amounts, benefiting gifted students, rather than compensating for the lack of family capital of poor rural youth.

Based on this analysis, I infer that innovation in Kazakhstan's education system in the first two decades of the post-Soviet period was sporadic, highly dependent on individual choices and interests of teachers and principals, and uneven, as higher ability schools tended to explore student-centered instruction to a larger extent than their peers from mainstream schools. Given the substantial variation in the understanding of what constitutes effective instruction, and the complete void of leadership in instructional reform have led the state to rely on external incentives and regulatory mechanisms for inducing quality. It appears evident that the lack of a conclusive answer to the question, "do student-centered or traditional teaching methods result in better achievement outcomes?" will not be resolved until a comprehensive, coherent strategy for instructional effectiveness is implemented by the state, and until a strong and sustainable system is built for continuous teacher quality improvement from within.

CHAPTER 8.
CONCLUSIONS AND AGENDA FOR FUTURE RESEARCH

8.1. OVERVIEW

In this study, I analyzed the relationship between student-centered instruction and achievement in mathematics and science in Kazakhstan, and examined the context of instructional environment in this post-Soviet country, with an added focus on the role of the state and its capacity to carry out on the vision it outlined for the country's educational development. To address the research questions outlined in Chapter 1, I employed a mixed methods approach, using both extensive quantitative analysis and a focused qualitative case study. The quantitative data from an international student achievement study, TIMSS 2007, were used to address the question of whether or not student-centered instruction was predictive of higher test scores of students in mathematics and science. The qualitative data collection served to shape my understanding of the context for the implementation of student-centered instruction, offering a more in-depth look at the perceptions and views of teachers, the challenges and incentives facing them, and the role and the relationship of the state vis-à-vis the schools.

8.1.1. MAIN RESEARCH HYPOTHESES

I based my analysis on a theoretical framework founded on two broad strands of literature: the literature on student-centered instruction and the learning theory behind the shift from traditional chalk-and-talk to child-centered pedagogy, on the one hand; and the literature on the role of the state in education, and specifically, on the role that the centralized state can play in creating an enabling environment for quality improvement in education, on the other. The Carnoy et al (2007) study prompted the construction of my

conceptual framework, providing a vivid example of extensive and successful state involvement in education, and offering a strong theoretical argument in favor of greater state role in building and sustaining quality of learning, at the expense of the individual choice of various actors. My theoretical framework, however, extends this argument by hypothesizing that change and innovation can be adopted through two broadly defined avenues, and state-led dissemination is one of them, while horizontal flow and adoption of successful practices is its alternative.

I constructed my research hypotheses at two levels. At the classroom level, I hypothesized that student-centered instruction, when it is implemented by a competent teacher, will generate a deeper and more intuitive understanding of the subject matter by the student, and consequently, will positively affect his or her achievement in that subject. At the system level, I hypothesized that in order to bring about results in terms of measurable student achievement outcomes, student-centered instruction must be part of the institutionalized, accepted forms of teaching, with a shared understanding among teachers and principals about how it is to be implemented. I further suggested that while there are (at least) two pathways to such an institutionalization – one through grassroots professional networks of teachers and the other through top-down, state-controlled comprehensive reform – it is far more likely that a centralized, authoritarian state will prefer the second approach to building quality, as it allows the state to solidify its grasp of the education system. Such an approach, however, is highly dependent on strong central capacity of the state, and hindered by the inability of the system to assimilate innovation and effective best practices introduced by non-state actors.

While the Carnoy argument is centered on the ability of the highly centralized state to restrict individual choices of actors in the education system in order to direct all efforts

towards a common goal, my researched showed that having the mechanisms of control at the disposal of the state alone is insufficient, and that it is important to acknowledge the specific types of leadership that are necessary for achieve quality in such a system. Furthermore, an increased level of centralization in decision making, as is the case in Kazakhstan, results in a system where innovation and quality improvement cannot take hold through horizontal dissemination, and the capacity of central government bodies to effect and sustain positive change determines the health of the entire system. Teachers lose (or, more precisely, never gain) the ability to act as agents of quality improvement, looking to the state to provide instructional leadership.

8.1.2. QUANTITATIVE RESULTS

The analysis of quantitative data showed no evidence that student-centered instruction played a role in student achievement in math and science. In my initial regional analysis, results across seven post-Soviet states revealed no clear pattern of relationship between student-centered methods and student test scores in these two cognitive domains, and showed that Kazakhstan is generally not substantively different in this regard from the other Eurasian states. In fact, it was the two traditional methods of instruction, memorization and independent work on solving science problems, that showed the largest positive point estimates in Kazakhstan (but not in the other states), compared to student-centered methods. In addition, when the variables measuring the frequency of exposure to a set of student-centered methods were combined to form a single measure of learning environment, the point estimates became negligible. Furthermore, through extensive modeling of student likelihood of being exposed to student-centered instruction in Kazakhstan, I found that while the probability of being in a student-centered environment was not completely random, there were no obvious, overwhelming biases in selection for

treatment that could have dramatically affected the results, neither in student background, nor in teacher and school characteristics. Indeed, with the most extensive bias control through propensity score matching, and control for the structure of the residual variance through multilevel modeling, point estimates were very small and statistically insignificant.

These findings led me to conclude that as of the time of the TIMSS assessment in 2007, student-centered practices had not affected student achievement outcomes in math and science. Furthermore, there was no evidence that the likelihood of the students' exposure to student-centered methods of instruction was determined by their background or family social capital. However, there is still a caveat to these conclusions: because there was no baseline assessment, and the measured scores are not gain scores, some unmeasured personal characteristics of students and their family background may have clouded the estimates of the benefit – or detriment – of student-centered instruction for their achievement scores. All of the observed characteristics of student background were controlled; but variables such as the education level of the parents and the extent to which the parents were involved in monitoring their children's academic performance, for example, remained unmeasured, leaving the researcher to wonder whether these important factors were sufficiently captured by the observed characteristics, such as the number of books in the student's home, or were left entirely out of the estimation process.

8.1.3. QUALITATIVE RESULTS

As I proceeded to the analysis of qualitative data, I found that while student-centered instruction was indeed the buzz word in policy talk, and teachers and principals largely supported the shift to a student-centered learning environment, there was no evidence that such a shift was taking place, or even that the state was envisioning for such a

shift to occur in the near future. Teachers were interested and intrigued by the new methodologies they had heard about from colleagues, but they had few ways to obtain first-hand experience in teaching using such methods, and even less so, to receive ongoing support and guidance in teaching their lessons. Support and guidance was especially emphasized given the reportedly low competence of newcomers into the profession, as a result of prolonged and continued neglect of state pre-service teacher training, and the relatively low starting wages within the sector, leading to a lower quality pool of applicants prior to the start of the training.

In addition to the lack of training, teachers were also simply overloaded with class hours, taking more than one teaching load (many took at least two), and pressed to continuously fill out paperwork for the district's routine data collection. Incentives were mostly set to reward rote learning methods, particularly in senior grades, with the exception of a small percentage of students who were academically gifted to be participating in academic Olympiads, thereby raising the prestige of their schools and teachers. However, the interest of the state in reviving and improving the condition of the education sector left no doubt, given the substantial increase of state funding for school infrastructure, salaries, educational technology, school feeding programs, and specialized and gifted schools – albeit,as the list shows, this interest was generally limited to highly visible, immediately traceable inputs, rather than the more tedious investment in instructional capacity at the schools and teacher training institutes. At the same time, teachers and instructors of training programs acknowledged the fact that as long as they complied with the state standards regimenting the number of class hours to be dedicated to subject matter content, they were free to explore their own methodologies for instruction, and were uninhibited in trying student-centered methods if they saw them fit. This freedom, however, often left teachers uncertain of what was expected of them, and they

continued to look to the state to provide leadership in building effective classroom instruction.

8.2. CONCLUSION: DOES STUDENT-CENTERED INSTRUCTION WORK FOR KAZAKHSTAN?

Going back to the theoretical framework laid out in Chapter 4, one is left with the question of whether the argument of the effectiveness of a centralized state in improving education, built so convincingly by Carnoy et al. (2007) on the Cuban case, is applicable to understanding the development of student-centered instruction in Kazakhstan. The answer to this question is multi-layered, and often not as clear-cut as one would like. The following main conclusions would form the basis of the argument.

8.2.1. PRESERVE AND CONTROL, OR LEAD AND ADVANCE? THE ROLE OF THE STATE IN KAZAKHSTAN'S EDUCATION SYSTEM

First, it is obvious that while in Kazakhstan, a shift to student-centered instruction was hailed as the most needed change for the education system, supposedly to make it more in-tune with the needs of the modern day social norms, the new civic identity of Kazakhstan's citizens in the 21st century, and the demands of the modern labor market for creativity and innovation, there was a substantial disconnect between the rhetoric of change in instruction and the actual dynamics of the state's interaction with the classroom. Unlike Cuba, Kazakhstan lacked a common narrative, a "moral imperative" that could bring all elements of a puzzle into a coherent whole, where a common understanding could be formed about what effective instruction means (traditional or student-centered), how to translate it into classroom activities, and what support teachers would require to ensure a consistent quality of implementation across the entire system. This lack of a common narrative is not, however, due to any structural limitations, or insufficient control of the

system. Much like in Cuba, the education system in Kazakhstan is centrally administered, and its regional and district-level education authorities are mere representative offices of the central Ministry of Education. Education policies and procedures, rules and regulations from the top to the very bottom, are developed and disseminated by the central state agencies. Reporting structures are similarly hierarchically organized, with central education authorities serving as the end point for information gathering, analysis, and administrative decision-making. The proximity of the state to the realities of teachers and education administrators was evident in interviews, through the respondents' references to overall education policy and to the official rhetoric of the state. The general anxiousness of interviewees about saying something on the record that might contradict the official state rhetoric also signaled the acute awareness of teachers of the authority and control mechanisms available to the state. Indeed, Kazakhstan possessed all of the advantages of centralized systems that Cuba used to its benefit: first and foremost, a direct reach into each component of the education system, and – in the absence of civic and professional organizations that would serve as a counterweight to its power – the ability to restrict individual choices if this was deemed necessary.

On the other hand, however, a key condition I proposed as being necessary for a successful state-driven, top-down institutionalization of student-centered teaching practices was not met in Kazakhstan: strong central capacity for leadership in instruction. Therefore, as much as the state had the *mechanisms* to introduce and disseminate a particular set of instructional approaches – a highly centralized network of in-service training institutes, a centralized curriculum and textbook development authority, and state-controlled pre-service teacher training programs – it was unable, and perhaps unwilling, to engage these mechanisms to serve a common vision. The state, in fact, encouraged teachers to seek out professional development opportunities outside of the state structures, thereby

creating a blend of state and market-based system for teacher training. This lack of capacity – or, perhaps, lack of genuine political will – to provide leadership in instruction resulted in a paradoxical situation, where teachers were severely restricted and extensively controlled through various curricular requirements and accountability and reporting mechanisms, but at the same time, were allowed substantial liberty in methods of instruction. This liberty, in turn, was translated into heterogeneity of classroom activities, and in the words of the one of the interviewees, a former school principal, into a situation where "everyone does as they please", and individual teachers and pre-service instructors relied on their personal, rather general understanding of how traditional methods were to be altered. Consequently, highly effective teachers with a strong interest in student-centered instruction were able to find ways to obtain training and classroom support through NGOs such as the Open Society Institute with its Step-by-Step methodology, while the majority were left unsure of what was expected of them, and whether they should seek out third party support or continue to wait for directives from the state.

As would be rational for a centralized state – an authoritarian state, in both Cuba and Kazakhstan cases – given its lack of leadership and management capacity, the efforts of the state were directed towards *protecting the status quo* in the education sector, particularly in instructional models, by pouring resources into existing structures and preventing outside parties such as NGOs from gaining too much influence, until the state itself becomes sufficiently strong to be able to regain control and implement its vision for educational development. Opposed to the idea of strengthening civic and professional associations that may challenge its authority, the state resists the institutionalization of teaching approaches at the grassroots levels, and reinforces the need to channel all innovation through its own mechanisms of identification, evaluation, and dissemination. At the same time, unable to fulfill the immediate needs of teachers for guidance and support,

and to provide leadership in instructional development, the state finds itself leaning on the free market, and putting the responsibility for teacher capacity building on the teachers themselves. This was reflected in the seemingly contradictory dynamics of NGO and private sector engagement with the state: educational NGOs were met with a lukewarm response by central authorities, and yet at the same time, the new education strategy included provisions for teachers to obtain vouchers for professional development outside the state in-service training system, effectively turning them back to the NGOs and other providers, and relying on the assumption of their own ability to distinguish good quality training courses from subpar ones. Strategies for improving the quality of pre-service teacher training were also absent, largely replaced by plans to restrict admissions, raise starting salaries for teachers, and conduct regular teacher re-certification, based on evaluations of their competence and professional fit – all of which seem to suggest a quest for individuals with intrinsic qualities of an effective teacher. Generally, in the sphere of teacher training, Kazakhstan shows a deviation from the Cuba's strong state model, offering a combination of state and market sources and suppliers (e.g. professional development vouchers).

In applying the framework of a strong state control of all elements of the system – based on Carnoy et al (2007) – to the case of Kazakhstan, it is important to recognize its limitations in understanding the factors that enable the success of a largely structural arrangement – such as centralized decision-making and resource planning – for boosting instructional quality. Steiner-Khamsi, Johnson, and Silova (2006) describe the clash between the policy reform rhetoric adopted by the highest levels of the state educational apparatus in Kazakhstan, and the underlying belief systems of educators at all levels of the national system, resulting in resistance to change. Discursive politics played center stage in the development of policy documents in the mid-2000's, with rhetoric on the importance of adopting policies seen as "modern" and "European" used to push reforms that were

originally seen as unnecessary meddling with an already well functioning education system. In that instance, discursive adoption of outcomes-based education (OBE) resulted in the acceptance of formal plans to introduce a twelve-year general education structure. In the years that followed, the lack of progress on OBE, along with slow but steady progress on the twelve-year system provide an indication that a lack of movement on a particular policy agenda item – such as student-centered instruction – may signal a genuine belief that no change or reform is required in that aspect of education. If this is true, one wonders whether the language of adoption of student-centered instruction serves a different purpose, advancing the agenda on policy priorities in which the state truly believes.

In sum, the case of Kazakhstan shows that having a firm grasp and tight control of the education system, as well as the ability to enforce state-driven policy comes with a challenge: the substantial demand on the state to possess the technical, intellectual, and managerial capacity to both generate a vision for the country's educational development, and see through its implementation at various levels. In undemocratic societies such Kazakhstan, the burden for leadership in quality improvement falls entirely on the state, given the absence of professional and civic organizations that could in other circumstances serve as breeding grounds for innovative ideas and exchange forums on effective practices. If such capacity is missing, there is no other avenue for the state but to solidify its control of the system through maintaining the status quo. Whether or not it eventually develops the capacity to lead may determine the ability of the system to sustain and improve the quality of student learning over the long term.

8.2.2. STUDENT-CENTERED INSTRUCTION: THE DEVIL IS IN THE DETAILS

This conclusion does not, however, complete our understanding of the impact and role of student-centered instruction in Kazakhstan. The heterogeneity in implementation, lack of a coherent common understanding of what student-centered instruction is among teachers, and most importantly, the lack of capacity on the part of the central state to engage the mechanisms available to it to implement its proclaimed vision make it difficult to expect positive results of student-centered teaching. Yet the absence of a substantive association with achievement cannot be fully explained away by these contextual factors alone. Fundamentally, these results lead to a different but related question: are student-centered approaches to teaching the way to go in educational development in Kazakhstan? Can we expect that with the Cuban level of effectiveness and synergy across elements of the education system, shifting to student-centered instruction would be the correct strategy to educate the new generations of citizens in this post-Soviet nation?

This question is left largely open by this dissertation. Results so far do not show a substantive difference in student test scores in math and science between the classrooms where student-centered instructional methods were practiced more frequently, and those where instruction was predominantly traditional chalk-and-talk. It is certainly true that the measures of student-centered instruction were somewhat noisy (fourth-grade student responses aggregated to classroom level), and could not capture the quality of student-teacher interactions – which are arguably as crucial to a truly student-centered environment as the quantity of such interactions. Furthermore, there is no pre-test measure, and therefore no way of controlling for prior achievement of the students. However, the complete lack of a somewhat substantive relationship between the overall

approach and achievement does give us pause, and calls for some reflection about what outcomes can be expected as a result of such a "paradigm shift" in instruction.

One thing is clear, both from the quantitative analysis and even more so, from teacher interviews: creating and measuring a student-centered learning environment is a formidable challenge in and of itself, even before any impact on achievement is contemplated. For the teacher, "stepping back" and letting the student be in the center of the learning process, while guiding and facilitating the learning activity to ensure that the student stays focused and learns important pieces of information is no easy task. Quite the opposite: it requires the teacher to possess not only superb knowledge of their subject matter, but also good judgment, an ability to manage individual and group-level tasks, and a capacity to ensure that no one is left behind in the process of individualized learning. Using such a methodology effectively puts even higher demands on the quality of teacher training and in-service support than the traditional lecturing or "cold-calling". In interviews conducted for this study, teachers lamented the lack of clarity in how to deliver their lessons, reminiscing about the "old days" when they were following scripted lesson plans. Implementing student-centered instruction effectively may require an even greater reliance on clear instructional support materials – perhaps also in the form of scripted lessons – until teachers develop a level of understanding and confidence that would allow them to use a broad range of teaching methods effectively without a script. Poor use of teaching techniques such as group work, for example, may result not in collective learning, but in instructional time loss, if students lose focus during a group activity, or if group dynamics result in some students being excluded from discussions. The criticisms made by some respondents in this dissertation included the perceived shift of all responsibility for learning away from the teacher as a result of student-centered activity, reflecting an understanding by some teachers that student-centered methods represent a way for the

teachers to have an easier time in class compared to a traditional instruction – when quite the opposite is often true.

The heterogeneity in the ways that teachers use student-centered methods, and the lack of clarity on what types of skills these methods require on the part of the teachers likely explain some portion of why no substantively significant (not just statistically significant) association could not be found in the TIMSS data. However, notwithstanding these factors, one must not rule out the possibility that even if they are implemented well and with equal effectiveness across the system, such methods may not, in fact, be as effective in improving student achievement in mathematics and science measured by standardized assessments as their proponents would hope. This explanation goes against my research hypothesis, in which I suggested that student-centered instructional methods, if implemented effectively, generate a deeper, more intuitive understanding of subject matter, which in turn manifests itself in higher achievement. The null hypothesis of no effect in this case cannot be rejected, which may mean that in fact these methods do not make a substantive difference on achievement, all other factors held constant. Perhaps it is indeed true that traditional instructional methods are more effective in *this* aspect of learning – cognitive achievement in math and science – although to what extent it *is* true is, of course, a subject of a different study and cannot be concluded from this dissertation. What we did see, however, in the course of the quantitative analysis, is that the two traditional methods of instruction that were used as "controls" in regressions of math and science scores on the measures of instructional environment, showed very large and substantive coefficients in the Kazakhstan's data, and therefore, are worthy of more in-depth research and analysis than was possible within the limits of this project. Furthermore, these initial findings suggest that perhaps different expectations of outcomes are appropriate for the different instructional approaches.

Interviews with teachers revealed that faced with the pressure of standardized assessments, they often reverted to the most basic forms of rote learning, asking students to memorize large amounts of information and retain it over the short-term period that covers the test. It is yet to be revealed, however, what portion of that information the students will retain after the tests are completed – and whether there is a difference in that deeper level of retention that can be attributed to one or the other instructional approach. This retraction to the teachers' comfort zone, where they are teaching to the test, and more generally – teaching in the ways they were themselves taught – is quite rational; and perhaps with more time and more experience with student-centered instruction, teachers would be just as comfortable – and just as effective – in generating test-score gains using child-centered pedagogy. However, because no such effects could be demonstrated to date (not just in Kazakhstan, but elsewhere in studies using TIMSS and PISA achievement data), one cannot argue against the use of more traditional methods, especially when it comes to measurable achievement results. If cognitive achievement is the goal above other skills, the evidence is still on the side of direct instruction. For this reason, reducing the emphasis from the teaching method per se, and putting it on teaching capacity in general – with training, induction, and in-service instructional support and resources – appears to be a more relevant strategy for short and medium term.

At this time, despite the rhetoric and vision statements about the importance of reform in education, the efforts of the state have been serving to reinforce old models and preserve the structures that were created during the Soviet period. The restoration of the physical plant, as well as highly visible inputs such as technology were by all means necessary and welcome by educators. However, a real engagement of the state in instruction and teacher capacity building has been lacking, and appears to be the single most important gap facing the education sector in Kazakhstan. While evidence is

inconclusive on whether or not student-centered instruction should replace traditional methods, it goes without question that the education system in the country is in serious need of instructional leadership on the part of the state, and requires a long-term, in-depth commitment to building a cadre of competent, effective teachers to replace the aging teacher corps trained predominantly in the Soviet era. This dissertation also shows that having strong central control over all parts of the system alone is insufficient – and it is specifically, instructional leadership and ability to affect the belief system of teachers and school administrators that is crucial in efforts to build or sustain instructional quality. Furthermore, a lack of political will and capacity at the central level of the education system to engage in instructional improvement, coupled with policies preventing the system from assimilating effective practices from outside (i.e. brought by third parties such as NGO's), makes the gap between high-level rhetoric and action all the more apparent.

The findings of this dissertation should also serve as a cautionary tale against linking a country's mean score on an international test immediately with that country's educational policy. It is true that Kazakhstan's test score on TIMSS was among the highest among all TIMSS participants in 2007; however, there is no evidence that any of the state's current policies had anything to do with this outcome. The education sector is much better resourced, and as a result of salary increases, teacher attrition is reportedly declined substantially over the last decade. However, as this dissertation demonstrates, no notable changes occurred at the classroom level; if anything, older teachers were unsure of what methods they were to follow, while younger teachers were trained in programs that had relied on outdated literature and were unprepared for the demands of classroom management. Given the substantially more modest results of the 2009 PISA assessment in Kazakhstan, it is fair to say that observers should always be careful about making causal linkages between state policy and classroom-level outcomes.

8.3. AGENDA FOR FUTURE RESEARCH

As I noted above, there are several areas where more research would be necessary in order to fully understand how, if at all, student-centered instructional methods affect achievement outcomes, and how the state can effectively improve instruction in general.

First, more studies such as this dissertation are necessary to form a core body of knowledge on the impact of student-centered instruction on achievement in different cognitive areas, grade levels of the students, and different student subpopulations. Specifically, it is important to know if some subject areas benefit from the use of student-centered instruction to a greater extent than others – such as, for example, mathematics and science. Similarly, it is also worth researching if younger students benefit more from student-centered instruction than older students – or vice versa, all other factors held constant. In this dissertation, I focused on the test scores of fourth graders in post-Soviet states. It is possible that results would have been different if the students were younger or older than the median age of ten in this sample.

Secondly, more studies are necessary to capture the *quality* of teacher-student interactions along with the *quantity* of such interactions – provided that such studies maintain the objectivity and random sampling of the large-scale studies such as TIMSS. Too often classroom observation data is limited to small sample projects, where the researchers are part of the implementation team, making it difficult to ascertain the extent to which the observations truly represent the population about which the researcher is making inferences (e.g., Marzano 1998). Nationally representative samples possess the benefit of direct generalizability of outcomes to the national population in question, and are therefore, in my opinion, more compelling and informative for education policy making.

Third, more evidence is needed on the impact of teacher-centered methods and the benefits and disadvantages of traditional instruction, referred to as "rote learning". In particular, any advantages of such methods over student-centered methods may be unduly overlooked in studies focusing on the positive aspects child-centered pedagogy, and therefore must themselves be a subject of a series of research investigations. The education community and especially teachers would benefit immensely from a deeper understanding of how direct instruction, such as lecture or memorization techniques, are necessary, and how they could be combined with other methods to create a student-centered yet productive learning environment.

Finally, returning to the level of the state and its role in educational transformation, more research is required to understand the dynamics of educational development in highly centralized, authoritarian environments. Several highly centralized education systems have shown quite impressive results in student achievement studies, but that alone is not sufficient to make any causal connections: there are almost as many top performers that are not authoritarian states. This dissertation also showed that one cannot make a direct connection between the student achievement outcomes and state policy. Further, the Carnoy et al (2007) argument about the advantages of a strong role of the state in education, often at the expense of individual choice and freedom, remains valid, but as my dissertation shows, the presence of strong control mechanisms alone is insufficient, and there are also substantial disadvantages and risks associated with complete concentration of control for the education system in the hands of the central state. It is important to distinguish between the types of state control: financial, administrative, and instructional – and it is precisely instructional leadership latter element of the role played by the state in education that is crucial for quality improvement. While change and innovation may be superficially present in policy documents, they remain at the discursive level – while the

status quo is continuously strengthened – until the state accumulates both the genuine political will and technical capacity to effect quality improvement. There are all indications that this discursive dynamic is taking place in Kazakhstan. With more research into the dynamics of centralized education system, their actors, their underlying incentives and ultimate goals, successes and failures, a richer understanding will be formed around the politics and policymaking in education in transitional countries such as Kazakhstan.

REFERENCES

Ammermuller, A., Heijke, H., and Woessmann, L. (2005) Schooling quality in Eastern Europe: Educational production during transition. *Economics of Education Review, 24* (2005), 579-599.

Antikainen, A. (1990). The rise and change of comparative planning: The Finnish experience. *European Journal of Education, 25*(1), 75-82.

Aypay, A., Erdogan, M. and Sozer, M. (2007). Variation among schools among classroom performances in science based on TIMSS-99 in Turkey. *Journal of Research in Science Teaching, 44* (10), 1417-1435.

Baker, D., and LeTendre, G. (2005). *National differences, global similarities*. Stanford, CA: Stanford University Press.

Ball, D.L., Lewis, J. & Thames, M.H. (2008). Making mathematics work in school. *Journal for Research in Mathematics Education, Monograph 14, A Study of Teaching: Multiple Lenses, Multiple Views*.

Bjork, C., and Tsuneyoshi, R. (2005). Education reform in Japan: Competing visions for the future. *The Phi Delta Kappan, 86*(8), 619-626.

Bransford, J., Brown, A., and Cocking, R., (2004). *How people learn: Brain, mind, experience, and school.* Expanded edition. Washington, DC: National Academy of Press.

Bryk, A., Raudenbush, S., and Congdon, R. (2010). *HLM 6.08 for Windows.* Lincolnwood, IL: Scientific Software International.

Carnoy, M., & Levin, H. M. (1985). *Schooling and work in the democratic state*. Stanford, CA: Stanford University Press.

Carnoy, M., with Gove, A., and Marshall, J. (2007). *Cuba's academic advantage: Why students in Cuba do better in school*. Stanford, CA: Stanford University Press.

Cave, P. (2001). Education reform in Japan in the 1990's: 'Individuality and other uncertainties. *Comparative Education 37(*2), 173-191.

Chan, D., and Mok, K. (2001). Educational reforms and coping strategies under the tidal wave of marketisation: A comparative study of Hong Kong and the mainland. *Comparative Education 37*(1), 21-41.

Chapman, D., Weidman, J., Cohen, M., and Mercer, M. (2005). The search for quality: A five-country study of educational strategies to improve educational quality in Central Asia. *International Journal of Educational Development 25* (5), 514–530.

Common, D. (1994). Conversation as a pedagogy of reform for public education. *The Journal of General Education, 43* (4), 241-272.

Creswell, J. (2008). *Research design: qualitative, quantitative, and mixed methods approaches.* Thousand Oaks, CA: SAGE.

Dabrowski, M. (1995). *Western aid conditionality and the post-Communist transition*. Center for Social and Economic Research. Warsaw, Poland: CASE. URL: http://ssrn.com/abstract=1479560 .

D'Agostino, R., and Rubin, D. (2000). Estimating and using propensity scores with partially missing data. *Journal of the American Statistical Association, 95* (451), 749-759.

Darling-Hammond, L. (2006). *Powerful teacher education: Lessons from exemplary programs.* San Francisco: Jossey-Bass.

Darling-Hammond, L., Barron, B., Pearson, P.D., Schoenfield, A.H., Stage, E.K., Zimmerman, T.D., Cervetti, G.N., and Tilson, J. (2008). *Powerful learning: What we know about teaching for understanding.* San Francisco: Jossey-Bass.

Davis, P., and Dombrowksi, P. (2000). International assistance to the former Soviet Union: Conditions and transitions. *Policy Studies Journal*, 28 (1), 68-95.

DeYoung, A., and Suzhikova, B. (1996). *Issues in post-Soviet secondary school reform: The case of Kazakhstan.* Published on Education Resources Information Center (ERIC): ED 403 104.

Doolittle, P. (1995). *Understanding cooperative learning through Vygotsky's zone of proximal development.* Published on Education Resources Information Center (ERIC): ED 484 575.

Drummond, T., and DeYoung, A. (2004) Perspectives and problems in education reform in Kyrgyzstan: The case of National Scholarship Testing. In Heyneman, S. and DeYoung, A. (Eds.). *The challenge of education in Central Asia,* pp. 225-244. Greenwich, Connecticut: Information Age Publishing.

Eklof, B., Holmes, L.E., and Kaplan, V. (Eds.) (2005). *Educational reform in post-Soviet Russia: Legacies and prospects*. London and New York: Frank Cass.

Falk, B. (2006). *Teaching the way children learn.* New York: Teachers College Press.

Figueroa, J. (1963). Selection and differentiation in Soviet schools. In King, E. (Ed.), *Communist Education*, pp. 124-152. New York: Bobbs-Merrill Company, Inc.

Foy, P. and Olson, J.F. (Eds.). (2009). *TIMSS 2007 user guide for the international database*. Chestnut Hill, MA: TIMSS and PIRLS International Study Center, Boston College.

Fuchs, T., and Woessman, L. (2004). *What accounts for international differences in student performance? A re-examination using PISA data.* IZA Discussion Paper, 1287. Bonn, Germany: The Institute for the Study of Labor (IZA).

Fujita, H. (2000). Education reform and education politics in Japan. *The American Sociologist, Fall 2000,* 42-57.

Gelman, A., and Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* New York: Cambridge University Press.

Hill, J., Rubin, D., and Thomas, N. (2000). The design of the New York school choice scholarships program evaluation. In L. Bickman (Ed.), *Research design: Donald Campbell's legacy.* New York: SAGE.

Holtz, R. (2002). Makarenko and Dewey: Two views on overcoming life circumstances through education. *Journal of Correctional Education*, 53 (3), 116-119.

House, J.D. (2005). Classroom instruction and science achievement in Japan, Hong Kong, and Chinese Taipei: results from the TIMSS 1999 assessment. *International Journal of Instructional Media*, 32 (3), 295-306.

Houtsonen, J., Czaplicka, M., Lindblad, S., Sohlberg, P., and Sugrue, C. (2010). Welfare state restructuring in education and its national refractions: Finnish, Irish, and Swedish teachers' perceptions of current changes. *Current Sociology, 58*(4), 597-622.

Iganaki, K. (1992). Piagetian and post-Piagetian conceptions of development and their implications for science education in early childhood. *Early Childhood Research Quarterly, 7*, 115-133.

Isaacs, N. (1972). *A brief introduction to Piaget.* New York: Agathon Press Inc.

Johnson, D. W, Johnson, R. T, and Stanne, M. B. (2000*). Cooperative learning methods: A meta-analysis*. Retrieved July 28, 2009 from http://www.co-operation.org/pages/cl-methods.html.

Kalyvas, S. (1999) The decay and breakdown of Communist one-party systems. *Annual Review of Political Science, 2*, 323-343.

Kissane, C. (2005) History education in transit: Where to for Kazakhstan? *Comparative Education,* 41 (1), 45-69.

Kim, J., and Mueller, C. (1978*). Introduction to factor analysis: What it is and how to do it.* Quantitative Applications in the Social Sciences Series #13. Thousand Oaks, CA: Sage.

Laitin, D. (1998) *Identity in formation: the Russian-speaking populations in the near abroad.* Ithaca, NY: Cornell University Press.

Leuven, E., and Sianesi, B. (2003) PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.

Libman, A., and Vinokurov, E. (2010). *Regional integration and economic convergence in the post-Soviet space: Experience of the decade of growth*. Munich Personal Re-Pec Archive (MPRA) Paper No. 21594. Published online at http://mpra.ub.uni-muenchen.de/21594 .

Lillard, A. S. (2005). *Montessori: The science behind the genius*. New York: Oxford University Press.

Makarova, M., and Solomennikov, V. (2008). The Bologna process: Opinions and expectations. *Russian Education and Society, 50* (3), 84-90.

Mertaugh, M. (2004). Education in Central Asia, with particular reference to the Kyrgyz Republic. In Heyneman, S. and DeYoung, A. (Eds.). *The challenge of education in Central Asia,* pp. 153-180. Greenwich, Connecticut: Information Age Publishing.

Mok, K. (2003). Decentralization and marketization of education in Singapore: A case study of the school excellence model. *Journal of Educational Administration 41* (4/5), 348-366.

Morris, P., Kan, F., and Morris, E. (2000). Education, civic participation and identity: Continuity and change in Hong Kong. *Cambridge Journal of Education, 30* (2), 243-262.

Montagnero, J., and Maurice-Naville, D. (1997). *Piaget, or The advance of knowledge.* Mahwa, New Jersey: Lawrence Erlbaum Associates, Inc Publishers.

Mullis, I., and Martin, M. (2007). TIMSS in perspective: Lessons learned from IEA's four decades of international mathematics assessments. In Loveless, T. (Ed.) *Lessons learned: What international assessments tell us about math achievement,* pp. 9-36. Washington, DC: Brookings Institution Press.

Olson, J.F., Martin, M.O., and Mullis, I.V.S. (Eds.). (2008). *TIMSS 2007 technical report.* Chestnut Hill, MA: TIMSS and PIRLS International Study Center, Boston College.

OECD. (2009). *Top of the class: High performers in science in PISA 2006*. Retrieved July 28, 2009 from www.oecd.org/pisa.

OECD (2010). *Strong performers and successful reformers in education: Lessons from PISA for the United States*. Paris: OECD.

Omoeva, C. (2011). Active learning methods and student achievement in math and science. In Silova, I. (Ed.) *Globalization on the margins: Education and postsocialist transformations in Central Asia*. Charlotte, NC: Information Age Publishing.

Papadopoulou, E. (2008). The humanitarian pedagogy of Sukhomlinsky and the application of his ideas in preschool education. *International Views on Early Childhood Education, 2008,* 1-9. Downloaded at http://sokl.joensuu.fi/verkkojulkaisut/varhais/papadopoulou.pdf .

Pliksnys, A., Kopnicka, S., Hrynevich, L., and Palicarsky, C. (2009). *Transparency in education in Eastern Europe.* Ethics and Corruption in Education series. Paris: UNESCO/ IIEP. Downloaded on 3/8/2010 from www.iiep.unesco.org/fileadmin/user_upload/ Info Services Publications/ pdf/2009/Transparencyineducation.pdf

Raudenbush, S., and Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods.* Thousand Oaks, CA: SAGE.

Reilly, D. (1996). Lessons from Soviet education: The need for an education system with responsibility, authority, and courage. *Journal of Educational Thought, 30,* 239-61.

Rosenbaum, P., and Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika,* 70(1), 41-55.

Rosenbaum, P. and Rubin, D. (1984). Reducing bias in observational studies using subclassification on the propensity score. *The Journal of the American Statistical Association, 79* (387), 516-524.

Rubin, D. and Thomas, N. (1996). Matching using estimated propensity scores: Relating theory to practice. *Biometrics*, 52, 249-264.

Rubin, D. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology, 2*, 169-188.

Rubin, D. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine, 127,* 757-763.

Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher, 39* (2), 142–151.

Schmidt, W.H., McKnight, C.C., Valverde, G.A., Houang, R.T., and Wiley, D.E. (1997). *Many visions, many aims: A cross-national investigation of curricular intentions in school mathematics.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Schmidt., W.H., Tatto, M.T., Bankov, K., Blomeke, S., Cedillo, T., Cogan, L., Han, S., Houang, R., Feng, J.H, Paine, L., Santillan, M., and Schwille, J. (2007). *The preparation gap: Teacher education for middle school mathematics in six countries (MT21 report).* East Lansing, MI: Michigan State University.

Sharpe, L., and Gopinathan, S. (2002). After effectiveness: new directions in the Singapore school system? *Journal of Education Policy, 17*(2), 151-166.

Shevyrev, A. (2005). Rewriting the national past: new images of Russia in history textbooks of the 1990's. In Eklof, B., Holmes, L.E., and Kaplan, V. (Eds.) *Educational reform in post-Soviet Russia: Legacies and prospects*. London and New York: Frank Cass.

Shimoniak, W. (1970). *Communist education: Its history, philosophy and politics.* Chicago, New York, San Francisco: Rand McNally and Company.

Silova, I. (2004). Adopting the language of the new allies. In Steiner-Khamsi, G. (Ed.) *The global politics of educational borrowing and lending*, pp. 75-87. New York: Teachers College Press.

Silova, I. (2008). Championing open society: The education logic of the Soros Foundation network. In Silova, I., and Steiner-Khamsi, G. (Eds.) *How NGOs react: Globalization and education reform in the Caucasus, Central Asia and Mongolia*, pp. 43-80. Bloomfield, CT: Kumarian Press.

Smith, G., Law, V., Wilson, A., Bohr, A., and Allworth, E. (1998). *Nation-building in the post-Soviet borderlands: The politics of national identities*. Cambridge: Cambridge University Press.

Standing, E. M. (1962). *The Montessori revolution in education*. New York: Schoken Books.

Steiner-Khamsi, G. and Omoeva, C., with Van Keuren, C. and Shiotani, A. (2009). *Access and quality in education: is there a trade-off?* Background paper for USAID Education Strategy. Washington, DC: USAID.

Steiner-Khamsi, G., Silova, I. and Johnson, E. (2006). Neo-liberalism liberally applied: Educational policy borrowing in Central Asia. In J. Ozga, T. Popkewitz and T. Seddon (Eds.), *Education research and policy: Steering the knowledge-based economy*, pp. 217-245. New York, NY: Routledge.

Steiner-Khamsi, G., and Stolpe, I. (2006). *Educational import: Local encounters with global forces in Mongolia*. New York: Palgrave Macmillan.

Su, Y., Gelman, A., Hill, J., and Yajima, M. (forthcoming, 2010). Multiple imputation with diagnostics (mi) in R: Opening windows into the black box. *Journal of Statistical Software*.

UNICEF (2011, forthcoming). Teachers: A regional study on the recruitment, development, and salaries of teachers in the CEECIS region. Geneva: UNICEF.

Vygotsky. L. S. (1978) *Mind in society.* Cambridge, MA: Harvard University Press.

Waddington, M. (1963). Russian children at home and at school. In King, E. (Ed.), *Communist education*, pp. 55-77. New York: Bobbs-Merrill Company, Inc.

Webber, S. (2000). *School, reform, and society in the new Russia.* New York: St. Martin's Press.

Wooldridge, J. (2002). *Introduction to econometrics: A modern approach*. Mason, OH: Thomson South-Western.